

Stereoskopische Ladeflächenerkennung und Füllstandsmessung am Beispiel von Häckselgut

Von der Fakultät für Elektrotechnik, Informationstechnik und Physik
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung der Würde

eines Doktor-Ingenieurs (Dr.-Ing.)

genehmigte Dissertation

von: Diplom Ingenieur Fabian Graefe

aus: Köln

eingereicht am: 11. September 2007

mündliche Prüfung am: 3. Dezember 2007

Vorsitz:

Prof. Dr. Ing. Dr. hc. em Werner Leonard

Referentinnen oder Referenten:

Prof. Dr. Ing. Walter Schumacher

Prof. Dr. Ing. Dr. hc. Hans-Heinrich Harms

Prof. Dr. Ing. Raul Queiroz Feitosa

2007

Inhaltsverzeichnis

Symbolverzeichnis	X
Kurzfassung	XVII
Abstract	XVIII
1 Einleitung	1
2 Grundlagen der stereo-optischen Geometrie	4
2.1 Die Modellparameter einer Kamera	4
2.2 Das Abbildungsmodell einer idealen Lochkamera	5
2.3 Die mathematische Beschreibung des Lochkameramodells	6
2.4 Das Abbildungsmodell einer Digitalkamera	8
2.4.1 Die Abbildungsfehler einer Linse	9
2.4.2 Die radiale Verzeichnung	10
2.4.3 Die tangentielle Verzeichnung	12
2.4.4 Die Abbildung auf einen realen Bildsensor	13
2.4.5 Die Abbildung auf einen diskretisierenden Sensor	14
2.4.6 Farbdarstellung	16
2.4.7 Sättigung und Luminanz	17
2.5 Stereo-optische Systeme	18
2.6 Die Epipolargeometrie	18
2.7 Die Homographiematrix H	20
2.8 Epipolargeometrische Beziehungen in Kamerakoordinaten	21
2.8.1 Die Essentialmatrix E	21
2.8.2 Die Epipolarlinie l	23
3 Der Bildeinzug	25
3.1 Der Prozess des Bildeinzugs	25
3.2 Die Anforderungen an den Bildeinzug	27
3.2.1 Belichtungsverfahren	28
3.3 Die Steuerung der Kameraelektronik	30
3.3.1 Die Bildqualität	30
3.3.2 Ein Gütefunktional zur Bewertung der Bildqualität	31
3.3.3 Die Regelung der Bildqualität	32
3.3.4 Die Qualität der eingezogenen Bildpaare	33
3.4 Die Kalibrierung des Epipolarsystems	34
3.4.1 Vorstellung der unterschiedlichen Kalibrierverfahren	35

3.4.2	Kalibrierung nach Zhang	35
3.4.3	Abschätzung der Homographiematrix \mathbf{H}	36
3.4.4	Bestimmung der intrinsischen und extrinsischen Parameter	37
3.4.5	Modellierung der Verzeichnung und Optimierung der Gesamtlösung	38
3.4.6	Kameraparameter des Messsystems	38
3.5	Die Optimierung der eingezogenen Bilder	43
3.5.1	Rektifizierung	43
4	Segmentierung	49
4.1	Der Prozess der Segmentierung	50
4.2	Interessierende Bildbereiche	51
4.2.1	Identifikation der Ränder	53
4.2.2	Farberkennung bei variabler Beleuchtung	55
4.2.3	Die DeHough-Transformation	56
4.2.4	Berechnung des DeHough- Raumes	56
4.2.5	Optimierung der DeHough-Transformation	60
4.2.6	Das geometrische Modell	64
4.3	Die Extraktion des Schüttgutkegels	67
4.3.1	Methoden der Texturklassifizierung	67
4.3.2	Verwendete Features und Frequenzanalyse	68
5	Stereovision	70
5.1	Der Prozess der Stereovision	70
5.2	Informationsgehalt eines Bildbereiches	73
5.3	Bildvorverarbeitung	78
5.4	Die Lösung des Korrelationsproblems (Pointmatching)	81
5.4.1	Das Pointmatching für das Ladegut	82
5.4.2	Die Bedeutung der Rektifizierung	85
5.4.3	Pointmatching für den Rand	88
5.5	Rekonstruktion	90
5.5.1	Die Triangulation	90
5.5.2	Die erweiterte Triangulation	91
5.5.3	Triangulation rektifizierter Bildpunkte	92
6	Ableitung des Füllstandes	95
6.1	Der Prozess zur Ableitung des Füllstandes	95
6.2	Die Definition der Messgröße <i>Füllstand</i>	96
6.3	Die Bestimmung der Ebene F_0	98
6.4	Koordinatensystem des Laderaumes	101
6.5	Transformation der Koordinatensysteme	103
6.6	Resample	105
7	Das Messsystem	108
7.1	Die Fehlerquellen	109
7.2	Der Einfluss des Disparitätsfehlers	110
7.3	Der Einfluss des Abbildungsfehlers	112

7.4	Kompensierbare Fehler	115
7.5	Die Ergebnisse	116
8	Schlussbemerkungen und Ausblick	119
A	Projektive Geometrie	122
A.1	Geraden	122
A.2	Punkte	122
A.3	Beziehung von Geraden und Punkten	123
A.4	Dualitätsprinzip	123
A.5	Punkte im Unendlichen	123
A.6	Modell der projektiven Ebene	124
A.7	Projektive Transformation	124
B	Transformationen im dreidimensionalen Raum	126
B.1	Rotation	126
B.2	Translation	127
B.3	Skalierung	127
B.4	Allgemeine Transformation	128
C	Levenberg-Marquardt-Algorithmus	129
D	Homogene Matrizen	131
E	FILLED Systemkonfiguration	133
	Literatur	134

Bildverzeichnis

1.1	Maisernte	1
2.1	Relevante Koordinatensysteme	5
2.2	Prinzip einer Lochkamera	5
2.3	Abbildung einer idealen Kamera	7
2.4	Abbildung mit einer realen Kamera	8
2.5	Strahlengang durch eine reale Linse	9
2.6	Radiale Linsen-Verzeichnung	10
2.7	Prinzip der radialen Verzeichnung	11
2.8	Radiale Verzeichnung	11
2.9	Prinzip der tangentialen Verzeichnung	12
2.10	Tangentiale Verzeichnung	13
2.11	Modell des Sensorchips	15
2.12	Achsparalleles Stereosystem (links), konvergente Anordnung (rechts)	18
2.13	Allgemeine Epipolargeometrie	19
3.1	Flussdiagramm des Funktionsblockes Bildeinzug	26
3.2	Abbildung in zwei optischen Systemen	27
3.3	Sensorabtastung im Rolling-Shutter-Betrieb	29
3.4	Verzerrte Abbildung bewegter Objekte durch sequentielle Abtastung	29
3.5	Regelungsstruktur	32
3.6	Eingezogene Bilder	33
3.7	Farbkanal-Histogramme der eingezogenen Bilder	33
3.8	Kalibrierungsmuster	36
3.9	Radiale Verzerrung der linken Kamera	40
3.10	Radiale Verzerrung der rechten Kamera	40
3.11	Tangentiale Verzerrung der linken Kamera	41
3.12	Tangentiale Verzerrung der rechten Kamera	41
3.13	Gesamtverzerrung der linken Kamera	42
3.14	Gesamtverzerrung der rechten Kamera	42
3.15	Rotation eines Bildes	45
3.16	Geometrische Deutung der bilinearen Interpolation	45
3.17	Rektifizierungsmatrix für die rechte Kamera	47
3.18	Beispiel einer Stereoansicht und ihrer Rektifikation	48
4.1	Flussdiagramm des Funktionsblocks Segmentierung	50
4.2	Segmentierung der interessierenden Bildbereiche	52
4.3	Ergebnis der Farbfilterung	53

4.4	Ergebnis der Median-Filterung	54
4.5	Ergebniss der <i>Canny</i> Kantendetektion	54
4.6	Gerade in Polardarstellung	57
4.7	DeHough Raum	58
4.8	Analytisch bestimmter De Hough Raum	59
4.9	Begrenzung des Bildbereichs(links), ROI im neu eingezogenen Bild	61
4.10	Vorsteuerung der Scanbereiche	62
4.11	Der Laderaum verlässt den Bildbereich	64
4.12	Eingabebild in die DeHough-Transformation	65
4.13	DeHough-Raum für einen begrenzten Winkelbereich	66
5.1	Flussdiagramm der Stereovisionsanalyse des Ladegutes	71
5.2	Flussdiagramm der Stereovisionsanalyse des Randes	72
5.3	Systematische Betrachtung von Punktekongstellationen	74
5.4	Informationsgehalt für das Korrelationsproblem	75
5.5	Topologie einer Textur	75
5.6	hinreichender Informationsgehalt eines Bildausschnittes	76
5.7	Berechnung des Informationsgehaltes eines Bildbereiches	77
5.8	Signal einer Pixelzeile mit einem abgeschatteten Bereich	78
5.9	Normiertes Signal einer Pixelzeile	79
5.10	Bild nach Optimierung der Bildqualität	81
5.11	Rektifiziertes Stereobildpaar des Laderaumes	83
5.12	Matrix der normierten Korrelationskoeffizienten	83
5.13	Korrelationskoeffizient bei rotierten Bildern	85
5.14	Suchfenster für den korrespondierenden Punkt	87
5.15	Pointmatching für den Rand	89
5.16	Geometrie der Triangulation	90
5.17	Geometrie der erweiterten Triangulation	92
5.18	Geometrie der Triangulation bei rektifizierten Bildern	93
6.1	Flussdiagramm der Stereovisionsanalyse des Ladegutes	96
6.2	Definition der Messgröße <i>Füllstand</i>	97
6.3	Koordinatensystem des Laderaumes	102
6.4	Nicht uniform verteilte Messpunkte	105
6.5	XPand - Auswahl der zu verarbeitenden Punkte	106
6.6	Rektilinear verteilte Messpunkte	107
7.1	Rekonstruktion eines verzeichnet abgebildeten Punktes	110
7.2	Fehler des rekonstruierten Punktes	111
7.3	Fehlerhafte Abbildung eines rektifizierten Bildpunktes	113
7.4	Disparitätsfehler durch Rektifizierung	113
7.5	Disparitätsfehler durch Rektifizierung	115
7.6	Versuchsaufbau für die Verifikation des Messsystems	116
7.7	Schnitt durch den Referenzkörper	117
7.8	Fehler der rekonstruierten Z-Koordinate	117

A.1	Ein Modell der projektiven Ebene [24]	124
A.2	Die projektive Transformation eines Punktes [24]	125

Tabellenverzeichnis

2.1	Farbdarstellung mit dem RGB-Farbmodell	17
3.1	Intrinsische Parameter: Linke Kamera	38
3.2	Intrinsische Parameter: Rechte Kamera	39
3.3	Extrinsische Parameter	39
3.4	Rektifizierte Kameraparameter	44
6.1	Ebenenparameter	101
E.1	Parameter der erprobten Systemkonfiguration	133

Symbolverzeichnis

Koordinatensysteme

S_0	Weltkoordinatensystem
$[X_0, Y_0, Z_0]$	Achsen des Weltkoordinatensystems S_0
S_K	Kamerakordinatensystem
$[X_K, Y_K, Z_K]$	Achsen des Kamerakordinatensystems S_k
S_{KL}	Kamerakordinatensystem der linken Kamera
S_{KR}	Kamerakordinatensystem der rechten Kamera
S_c	Koordinatensystem der Bildebene (i.e. Kamera-Sensor)
S_{sc}	diskretes Koordinatensystem des Kamera-Sensors
S_R	Koordinatensystem des Laderaumes
$[u_c, v_c]$	Achsen des Koordinatensystems der Bildebene S_c
$[u_{sc}, v_{sc}]$	Diskrete Achsen des Koordinatensystems des Kamerasensors S_{sc}
\vec{e}_{u, S_c}	Einheitsvektor der u -Achse der Bildebene
\vec{e}_{v, S_c}	Einheitsvektor der v -Achse der Bildebene
\vec{n}_{S_c}	Normalen-Vektor auf die Bildebene
O_j	Ursprung des i -ten Koordinatensystems
$\vec{O}_j _{S_l}$	Ortsvektor zum Ursprung des i -ten Koordinatensystems bzgl. des l -ten Koordinatensystems
k_u, k_v	Anzahl der Pixel entlang der u - bzw. v -Achse des Sensors

Punkte und Vektoren

P	Allgemeiner Punkt/Objektpunkt im \mathbb{R}^3
$P _{S_j}$	Punkt im j -ten Koordinatensystem
P_{i_j}	i -te Komponente eines Punktes P im j -ten Koordinatensystem
P'	Verzerrt abgebildeter Punkt
\bar{P}	Unverzerrt abgebildeter Bildpunkt
\tilde{P}	Rektifizierter Punkt
P_p	Hauptpunkt oder Principal Point
P_L, P_R	Bildpunkt in der linken bzw. rechten Bildebene
P^*	Korrespondierender Punkt zu einem gewählten Referenzpunkt
$\delta P(P)$	Versatzvektor der Linsenverzeichnung als Funktion des Punktes P
\vec{P}_M	Vektor der markierten Pixel für die DeHough Transformation
P_{Mi}	i -tes, markiertes Pixel für die DeHough Transformation
\mathbf{P}_B	Vektor der Randpunkte
$\vec{R}_{P_{Mi}}$	Ortsvektor zum Punkt P_{Mi}
\vec{t}	Translationsvektor
\vec{t}_{s_j}	Translationsvektor zum j -ten Koordinatensystem
\vec{t}_i	i -te Komponente des Translationsvektors \vec{t}
\vec{n}_{F_0}	Normalen Vektor auf die F_0 -Ebene
\mathbf{t}_\times	Matrix zur Darstellung des Kreuzproduktes als ein Matrix-Vektor-Produkt
\vec{V}_P	Einheits-Richtungsvektor des Abbildungsstrahls des Punktes P
θ_i	Rotatorischer Versatz um die i -te Achse
\vec{R}	Rotationsvektor
\mathbf{R}	Rotationsmatrix
\mathbf{R}_{S_K}	Rotationsmatrix von Weltkoordinaten in Kamerakoordinaten

Segmentierung

q_B	Quotient aus Intensitätswert des Rot- und des Grün-Layers
q_G	Quotient aus Intensitätswert des Rot- und des Blau-Layers
T_B	Schwellwert für q_B zur Farbdefinition
T_G	Schwellwert für q_G zur Farbdefinition

Parameter der Abbildung

n, n_L, n_R	Bildebene, linke Bildebene, rechte Bildebene
m	Gleichmaßstäbigkeit der Sensorkoordinaten
s	Scherung der Sensorkoordinaten
\vec{c}	Vektor der Aufnahme­richtung
c	Kamerakonstante i.e. $ \vec{c} $
F	Brennpunkt
f	Brennweite
H_1, H_2	Hauptebenen einer Bikonvexlinse
K_1, K_2	Schnittpunkt der optischen Achse mit H_1 bzw. H_2
K_{Pz}	Projektionszentrum einer Linse
\mathbf{A}	Abbildungsmatrix einer Lochkamera
\mathbf{A}_2	Matrix der Abbildung auf einen diskretisierenden Sensor
\mathbf{A}_i	Abbildungsmatrix für die i -te Kamera
r	Radius eines Punktes bzgl. des Projektionszentrums einer Linse
\tilde{r}	Verzeichneter Radius eines Punktes bzgl. des Projektionszentrums einer Linse
k_1, \dots, k_i	Radiale Linsenverzerrungskonstanten
p_1, \dots, p_i	Tangentiale Linsenverzerrungskonstanten
B	Basislinie
\vec{V}_B	Basisvektor eines optischen Stereosystems
\mathbf{H}	Homographiematrix
$\mathbf{H}_{PP'}$	Homographiematrix für die Abbildung von P auf P'
$\hat{P}'(P_{ij}, \dots)$	mit den angegebenen Parametern reprojezierte Abbildung des Punktes P_{ij}
s	Vergrößerungsfaktor für die Abbildung durch eine Kamera
e_L, e_R	Epipol im linken bzw. rechten Bild
$l_{L,P}, l_{R,P}$	Epipolarlinie im linken bzw. rechten Bild zu einem Punkt P
g	allgemeine Gerade im \mathbb{R}^2
\vec{g}	Koeffizienten Vektor der Geraden g
h, i, j	Parameter einer allgemeinen Geraden im \mathbb{R}^2
π_P	Epipolarebene zu einem Punkt P
\mathbf{E}	Essentialmatrix
\mathbf{F}	Fundamentalmatrix
Z_i	Strukturparameter i.e. Entfernung eines Punktes P zum Kamerasystem
\mathbf{M}_K	Rektifikations-Matrix für die k -te Kamera

Bildeigenschaften und Bildeinzug

I	Allgemeines Bild
E_q	Regelabweichung des Bildqualitätswertes
E_{q0}	Schwellwert der E_q -Regelabweichung für die Steuerung der Initialisierung
H_i	Histogramm der Pixelwerte des i -ten Bildes eines Bilddatensatzes
HD_d	Resultierendes Histogramm aus der absoluten Differenz der Histogramme H_i
Q_i	Bildqualitätswert des i -ten Bildes eines Bilddatensatzes
Q_{soll}	Sollwert der Bildqualität
B_{mean}	mittlerer Helligkeitswert eines Bildes
B_{max}	maximaler Helligkeitswert eines Bildes
B_{clip}	Relativer Anteil an Clipping Pixeln
$H(255)$	Eintrag des 255ten Kontainers eines Histogramms
$H(255)_{max}$	Maximal zulässiger Eintrag des 255ten Kontainers eines Histogramms
I_i	Menge der auftretenden Pixel bzw. Intensitätswerte der i -ten Bildebene
a, b, c	Gewichtungsfaktoren zur Bemessung der Bildqualität Q
V_{CCD}	Verstärkung des CCD-Chips
t_{exp}	Exposuretime, Belichtungszeit einer Kamera
t_{lat}	Verzögerungszeit der Kamerahardware
k	Blendenwert

Stereovision

IC_{uv}	Informationsgehalt eines Bildausschnittes an der Stelle uv
I_{TP}	Tiefpass gefiltertes Bild
I_{HB}	<i>Highboost</i> gefiltertes Bild
h_{GTP}	Gauß-Tiefpassfilter
I_{lc}	Bild, das den lokalen Kontrast repräsentiert
I_n	Bild, das den normierten lokalen Kontrast repräsentiert
N_P	Bildausschnitt um den Punkt P in linken Bild
i, j	Größe des Bildausschnittes N_P
a_{ij}	Wert des Pixels an der Stelle i, j aus N_P des linken Bildes
b_{ij}	Wert des Pixels an der Stelle i, j aus N_P des rechten Bildes
N_R	Bildbereich im rechten Bild
k, l	Größe des Bildbereiches N_R im rechten Bild
C_M	Matrix der normierten Korrelationskoeffizienten

Stereovision

D_P	Disparität der Abbildungen eines Punktes P in einem Epipolarsystem
P_{Li}	$3 \times n$ Vektor des i -ten ausgewählten Punktes im linken Bild
P_{Ri}	$3 \times n$ Vektor des i -ten ausgewählten Punktes im rechten Bild
P_{Oi}	$3 \times n$ Vektor der rekonstruierten Objektpunkte
O_L	optisches Zentrum des linken Bildes
O_R	optisches Zentrum des rechten Bildes
\mathbf{E}	Essential Matrix
\mathbf{F}	Fundamental Matrix
\mathbf{E}_{Rect}	Essential Matrix des rektifizierten Bildes
S_{Li}	Strahlen vom Punkt P_{Li} durch das optische Zentrum der Kamera O_L
S_{Ri}	Strahlen vom Punkt P_{Ri} durch das optische Zentrum der Kamera O_R
\vec{V}_c	kürzester Verbindungsvektor zweier windschiefer Geraden
\vec{V}_{Li}	Richtungsvektor, der S_{Li} definiert
\vec{V}_{Ri}	Richtungsvektor, der S_{Ri} definiert
μ	Vielfaches des normierten Verbindungsvektors zwischen dem optischen Zentrum der rechten Kamera und dem Objektpunkt
λ	Vielfaches des normierten Verbindungsvektors zwischen dem optischen Zentrum der linken Kamera und dem Objektpunkt
γ	Vielfaches des kürzesten Verbindungsvektors zwischen den Abbildungsstrahlen
F_0	Ebene, die der obere Rand des Laderaumes definiert
$\varphi_{n_{F_0}}$	Winkelversatz der F_0 Ebene um die Z -Achse bezüglich des WKS
\mathbf{R}_{Rod}	Rodrigues Matrix zur Rotation der F_0 -Ebene
h_{xy}	Füllhöhe an der Stelle x,y im Laderaum
f_{xy}	inverse Füllhöhe an der Stelle x,y im Laderaum
$\vec{P}_{i,G}$	Der i -te Punkt auf der rekonstruierten Schüttgutoberfläche
$\vec{P}_{i,B}$	Der i -te Punkt auf dem rekonstruierten Rand
\vec{P}_G	$3 \times n$ Vektor der Punkte, die Element der rekonstruierten Schüttgutoberfläche sind
\vec{P}_B	$3 \times n$ Vektor der Punkte, die Elemente des rekonstruierten Randes sind
\vec{P}_m	$2 \times n$ Vektor der Pixel, die den Rand repräsentieren
\mathbb{R}^3	Dreidimensionaler Raum der Objektpunkte
\mathbb{R}^2	Zweidimensionaler Raum einer Ebene
\mathbb{P}_R^*	Menge der Kandidaten für korrespondierende Punkte

DeHough Transformation

A_{mi}	DeHough-Akkumulatorzelle für einen Bildpunkt P_{Mi}
$G_{P_{mi}}$	Menge der Geraden g durch einen Punkt P_{Mi} der Bildebene
ρ_{mij}	Polarparameter einer Geraden g_j durch den Punkt P_{Mi} im \mathbb{R}^2
φ_{mij}	Polarparameter einer Geraden g_j durch den Punkt P_{Mi}
$\delta\varphi$	Winkelauflösung der DeHough Transformation
$\rho(P_{mi}, \varphi)$	Abstand einer Geraden durch den Punkt P_{mi} mit dem Winkel φ zum Ursprung
$\vec{c}(\varphi)$	Einheits-Richtungsvektor einer Geraden als Funktion des Winkels φ
P_{mi}	im \mathbb{R}^2 i.e. Winkel des normalen Aufpunktvektors zur X -Achse
n	Indexzähler des Winkelauflösung für φ
n_{max}	Maxzahl der Winkeldiskretisierungsschritte für φ
n_P	Anzahl der Punkte Tupel
m	Anzahl der markierten Pixel
C_{gp}	Anzahl der auf einer Linie akkumulierten Pixel
$C_{gp,krit}$	Mindestanzahl der auf einer Linie akkumulierten Pixel für die Plausibilisierung der Existenz einer Geraden
\mathbf{D}_I	DeHough-Raum für ein Bild I
\vec{A}_g	Aufpunktvektor auf eine Gerade g
\mathbf{I}	Matrix die ein Bildrepräsentiert
$I_{\rho\varphi}$	Häufigkeit in der ein ρ/φ -Tupel auftritt

Fequenzanalyse

k_{uv}	Bildausschnitt um das Pixel u, v
pw_K	Leistungsspektrum der sichtbaren Texturen im Bildausschnitt k
\mathbf{F}	Matrix der Fourierkoeffizienten
f_{pq}	Fourierkoeffizient der p -, q -ten Frequenz aus \mathbf{F}'
T_0	Schwellwert des Fourierkoeffizienten für die Texturklassifizierung
T_1	Schwellwert des Fourierkoeffizienten für die Texturklassifizierung
\vec{S}_i	Vektor der Schnittpunkte der im Bild sichtbaren Geraden
u_B, v_B	Begrenzung des interessierenden Bereichs
r_F	Frequenzradius
$n_{\mathbb{F}}$	Anzahl der Elemente in \mathbb{F}

Füllstand

h_{xy}	lokale Füllhöhe des Ladegutes an der Stelle x, y des Laderaumes
F	Matrix, die den momentanen Füllstand repräsentiert
d_{krit}	Ausschlussentfernung für die Auswahl der Punkte für die Bestimmung der F_0 -Ebene
$d_{P_{Bi}}$	Abstand eines Randpunktes P_{Bi} zur Ebene F_0
n_i	i -ter Parameter einer Ebenengleichung in der Hesseschen Normalform
E	Allgemeine Ebene im \mathbb{R}^3
g_{ri}	Gerade die den i -ten Rand analytisch beschreibt

XPand-Algorithmus

r	Radius des Einzugsgebietes
H_r	Stelle des zu schätzenden Messwertes
h_{xy}	Gemessener Füllstabdnd an der stelle x, y
\mathbb{P}	Die Menge der berücksichtigten Punkte

Abkürzungen

WKS	Weltkoordinatensystem
KKS	Kamerakordinatensystem
SKS	Sensorkoodinatensystem
CCD	Charge coupled device
ROI	Region of Interest

Kurzfassung

Für die Automatisierung des Überladeprozesses zweier in Bewegung befindlicher Fahrzeuge sind zwei Informationen notwendig. Einerseits muss die relative Position der Fahrzeuge zueinander bekannt sein. Andererseits ist es wichtig, den Beladungspunkt auf der Ladefläche des Transportfahrzeuges zu kennen. Im Allgemeinen muss bei der Verladung von Schüttgütern der Beladungspunkt während der Beladung variiert werden, um eine optimale Ausnutzung des Laderaumes zu erreichen. Es ist sinnvoll, die Stellen, an denen eine vollständige Ausfüllung des Laderaumes noch nicht erreicht wurde, bevorzugt zu befüllen. Eine Voraussetzung für eine so *geregelte* Befüllung ist die Erfassung der Verteilung des Ladegutes im Laderaum. In diesem Umfeld entwickelt das Institut für Regelungstechnik der Technischen Universität in Braunschweig das System FILLED zur Videodaten basierten Füllstandserkennung.

Schlagwörter:

- Füllstandserkennung
- Bildverarbeitung
- Stereovision
- Überladeprozess
- automatisches Überladen

Abstract

For automation of a continuous overloading process between two vehicles in motion, two parameters are essential. On the one hand the relative position between the vehicles has to be known. On the other hand the loading point within the load space of the transport car has to be determined. Often only a non optimal usage of the transport capacity is obtained without moving the overload swivel. In order to optimize the filling by adjusting load point the distribution of the freight within the load space has to be measured during the overload process. In this article the Institut für Regelungstechnik of the Technische Universität in Braunschweig introduces the system FILLED for video data based fill level detection of agricultural bulk freight such as chaffed corn or grass.

Keywords:

- machine vision
- fill level detection
- stereo vision
- overload process
- automatic overloading

1 Einleitung

In der Landwirtschaft ist ein Trend zu immer stärkeren und größeren Erntemaschinen zu erkennen. Dies ermöglicht immer größere Erntegeschwindigkeiten. Mit diesem Anstieg der Erntegeschwindigkeit und Erntegutdurchsatz steigt gleichermaßen die Belastung für die Fahrer der Erntemaschinen. Die Abbildung 1.1 zeigt eine typische Überladungssituation aus der Maisernte.



Bild 1.1: Maisernte

In diesem Umfeld wurde an der Technischen Universität in Braunschweig in einer Kooperation des Instituts für Landmaschinen- und Fluidtechnik mit dem Institut für Regelungstechnik das System *ASÜL* [2, 3] zur Automatisierung des Überladevorgangs eines selbstfahrenden Feldhäckslers entwickelt. Mit dem Ziel, die Fahrer einer Erntekombination, bestehend aus

einem selbstfahrenden Feldhäcksler und einer Transporteinheit zu entlasten, soll das Assistenzsystem *ASÜL* automatisch den Überladepunkt im Bereich der Ladefläche positionieren. Die für die automatische Positionierung relevanten Informationen sind einerseits die relative Position der beteiligten Fahrzeuge zueinander und andererseits der Beladungspunkt innerhalb der Ladefläche.

Um hinsichtlich der optimalen Auslastung des Transportvolumens den momentan optimalen Beladungspunkt bestimmen zu können, ist es notwendig, die Verteilung des Erntegutes auf der Ladefläche zu kennen. In diesem Zusammenhang konnte im Rahmen dieser Arbeit am Institut für Regelungstechnik das System *FILLED* zur Messung des Beladungszustandes eines Schüttgutes in einer landwirtschaftlichen Transporteinheit entwickelt werden. *FILLED* erstellt stereoskopisch ein dreidimensionales Modell der Ladegutoberfläche sowie des Laderaumes selbst und leitet aus diesen Informationen die lokale Füllhöhe des Ladegutes ab. Das Messsystem *FILLED* wurde weitestgehend anwendungsnah spezifiziert und umgesetzt. Die wesentlichen Forderungen an dieses optische Messsystem, neben der Erfassung des Füllstandes selbst, war die Funktionsfähigkeit des unter den in der Mais- oder Grasernte üblichen Umgebungsbedingungen. Hierbei wurde besonderes Augenmerk auf die Robustheit gegen die natürlichen Beleuchtungsvarianzen, Schatteneinflüsse sowie der Relativbewegung der beteiligten Fahrzeuge gelegt.

Viele bestehende optische Messsysteme werden unter eindeutig definierten Einsatzbedingungen eingesetzt. Eine Herausforderung ist es, *FILLED* gegen die Vielfalt der möglichen Störeinflüsse natürlicher Beleuchtung unempfindlich zu gestalten. Darüber hinaus existiert eine Menge verschiedener Messgeräte zum Erfassen von Füllständen der unterschiedlichsten Speicherräume. Die besonderen Eigenschaften des hier betrachteten Ladegutes bedürfen einer Erfassung vieler über die Fläche des Laderaumes verteilter Messstellen, da sich das Ladegut nicht - wie bei Fluiden - gleichmäßig im Laderaum verteilt. Messeinrichtungen, die nur einen punktuellen Füllstandsmesswert liefern, waren für diese Aufgabe also nicht geeignet. Ultraschall basierte Messsysteme können hier aufgrund der rauen Umgebungsbedingungen nicht eingesetzt werden. Radarsensorik ist wegen ihrer Schwäche, organische Materie zu vermessen, nicht geeignet. Jegliche Sensorik, die an das Transportmittel gebunden ist, wurde aus ökonomischen Erwägungen und mit Hinsicht auf die Anwendbarkeit abgelehnt. Versuche mit Laserlichtlaufzeitmesssystemen haben gezeigt, dass die Dichte der Messstellen nicht ausreicht, um Störungen durch umherfliegende Partikel als solche zu detektieren, was eine Vermessung der Oberfläche des Ladegutes während der Beladung unmöglich machte.

Mit dem Messsystem *FILLED* kann somit ein neuartiges Messsystem zur passiv-optischen und berührungslosen Erfassung des Beladungszustandes einer landwirtschaftlichen Trans-

portereinheit vorgestellt werden. Mit *FILLED* ist es möglich, die Stellen im Laderaum zu erkennen, an denen eine vollständige Auslastung der Transportkapazität noch nicht erreicht wurde. Somit ist für das Assistenzsystem *ASÜL* das letzte Element, das für die vollständige Automatisierung des Überladungsvorgangs notwendig ist, geschaffen worden.

2 Grundlagen der stereo-optischen Geometrie

In diesem Abschnitt wird auf die Grundlagen der Erfassung von Bildern durch Ein- und Mehr-Kamerasysteme eingegangen. Ausgehend vom Abbildungsverhalten einer idealen Lochkamera wird die mathematische Formulierung des Abbildungsverhaltens von digitalisierenden Kameras erläutert. Darauf aufbauend werden die Zusammenhänge der stereooptischen Abbildung beschrieben.

2.1 Die Modellparameter einer Kamera

Die Modellparameter einer Kamera werden als extrinsische und intrinsische Kameraparameter bezeichnet. Die extrinsischen Kameraparameter beschreiben die Lage des Koordinatensystems S_K einer Kamera im Weltkoordinatensystem S_O . Die intrinsischen Parameter einer Kamera beschreiben die interne Geometrie einer Kamera sowie den Strahlengang durch deren Objektiv. Sie werden nicht von der Lage der Kamera im Weltkoordinatensystem beeinflusst. Die intrinsischen Kameraparameter beschreiben somit die Abbildung eines dreidimensionalen Objektpunktes als Bildpunkt auf der Bildebene.

Für die weiteren Betrachtungen werden entsprechend der Abbildung 2.1 die folgenden Koordinatensysteme eingeführt:

- Weltkoordinatensystem S_0 mit den Achsen $[X_0, Y_0, Z_0]$
- Kamerakoordinatensystem S_k mit den Achsen $[X_k, Y_k, Z_k]$
- Koordinatensystem der Bildebene S_c mit den Achsen $[u_c, v_c]$
- Koordinatensystem des Sensorchips S_{sc} mit den Achsen $[u_{sc}, v_{sc}]$

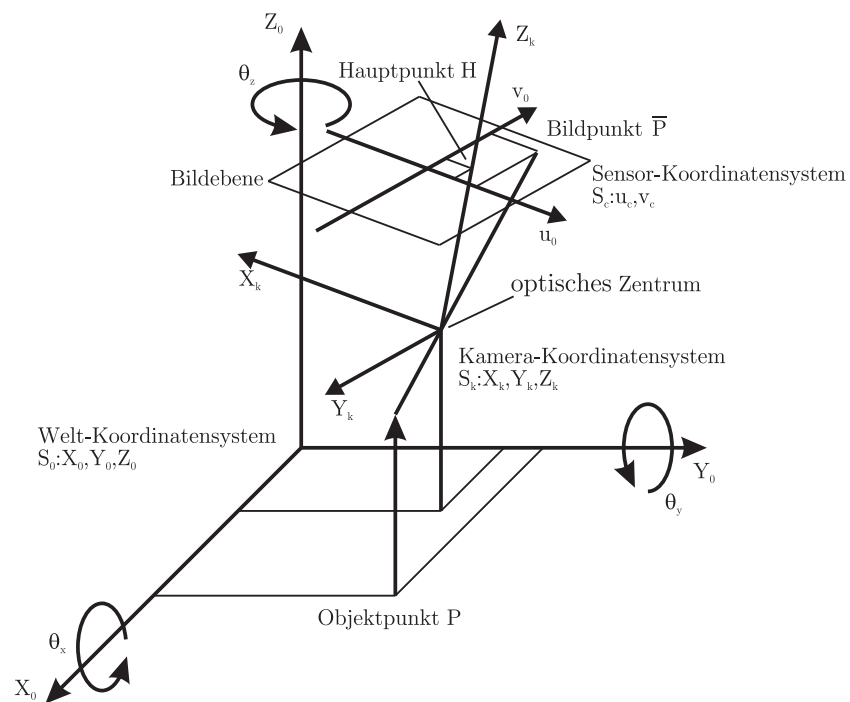


Bild 2.1: Relevante Koordinatensysteme

2.2 Das Abbildungsmodell einer idealen Lochkamera

Die ideale Lochkamera besteht aus einer Blende mit unendlich kleiner Blendenöffnung sowie einer Bildebene, deren Normalenvektor parallel zur Aufnahmerrichtung der Kamera liegt. Die Blende der Lochkamera stellt das perspektivische Zentrum des optischen Systems dar. Das perspektivische Zentrum wird auch als optisches Zentrum bezeichnet. Diese Kamera liefert durch die ideale perspektivische Abbildung auf der Bildebene unverzerrt abgebildete Bilder. Die Abbildung 2.2 zeigt das Prinzip der Lochkamera.

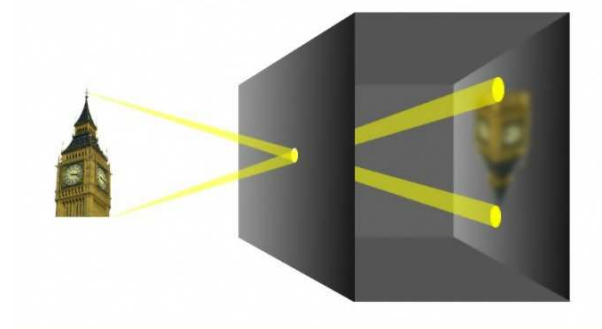


Bild 2.2: Prinzip einer Lochkamera

2.3 Die mathematische Beschreibung des Lochkameramodells

Die Transformation des Objektpunktes P von Welt- in Kamerakoordinaten wird durch den Translationsvektor \vec{t}_{S_k} und die Rotationsmatrix \mathbf{R}_{S_k} beschrieben. Die Gleichung 2.1 beschreibt die Transformation von Weltkoordinaten in Kamerakoordinaten.

$$P|_{S_k} = [\mathbf{R}_{S_k} \vec{t}_{S_k}] P|_{S_0} \quad (2.1)$$

Im Folgenden werden alle Punkte, wenn nicht anders definiert, in homogenen Koordinaten angegeben. Die Rotationsmatrix \mathbf{R}_{S_k} ergibt sich nach Gleichung 2.2. Hierbei bemessen die Winkel $\theta_x, \theta_y, \theta_z$ die Drehung der Kamera um ihre X -, Y -, Z -Achse bezüglich des Weltkoordinatensystems.

$$\mathbf{R}_{S_k} = \mathbf{R}_x \mathbf{R}_y \mathbf{R}_z \quad (2.2)$$

Die Matrizen \mathbf{R}_x , \mathbf{R}_y und \mathbf{R}_z ergeben sich nach

$$\mathbf{R}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{pmatrix} \quad (2.3)$$

$$\mathbf{R}_y = \begin{pmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{pmatrix} \quad (2.4)$$

$$\mathbf{R}_z = \begin{pmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.5)$$

Hierbei ist zu beachten, dass Matrizenmultiplikationen nicht kommutativ sind. Dementsprechend muss, wie in Gleichung 2.1 beschrieben, zunächst die Rotation und dann die Translation ausgeführt werden, um einen Punkt vom einen in das jeweils andere Koordinatensystem zu transformieren. Somit ist die Reihenfolge der Multiplikation in Gleichung 2.2 und 2.1 maßgeblich.

Die Abbildung des Punktes $P|_{S_K}$ (vgl. Abbildung 2.1) im Kamerakoordinatensystem S_K auf den Punkt $\bar{P}|_{S_c}$ im Koordinatensystem S_c der Bildebene durch eine Lochkamera beschreibt Gleichung 2.6.

$$\bar{P}|_{S_c} = \mathbf{A} P|_{S_K} \quad (2.6)$$

Hierbei ist \mathbf{A} eine 3×3 -Matrix und wird als Projektionsmatrix bezeichnet.

Die Abbildung 2.3 zeigt die Geometrie der Abbildung durch eine ideale Lochkamera.

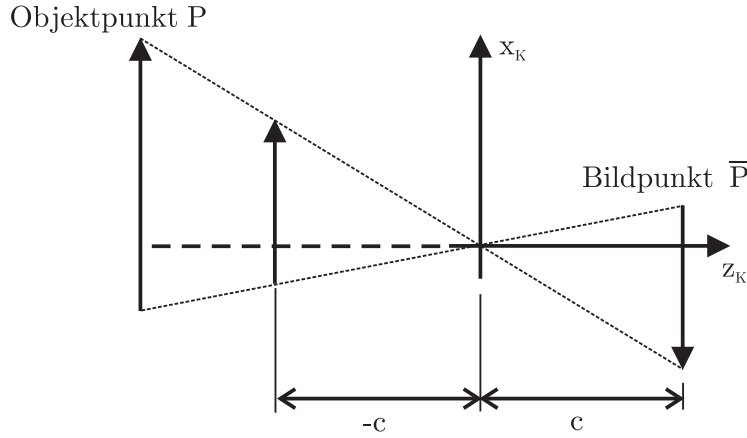


Bild 2.3: Abbildung einer idealen Kamera

Der Abstand zwischen der Bildebene und dem Projektionszentrum wird als Kamerakonstante c bezeichnet. Die Kamerakonstante ist gleich der Bildweite des optischen Systems. An Abbildung 2.3 wird deutlich, dass die Objekte am Projektionszentrum punktgespiegelt abgebildet werden, was für viele Anwendungen ungünstig ist. Verwendet man eine negative Kamerakonstante so wird das Objekt ungespiegelt abgebildet. Die Abbildung entspricht nun einer zentrischen Streckung. Somit ergibt sich die Abbildung eines Punktes P auf den Punkt \bar{P} nach Gleichung 2.7

$$\begin{bmatrix} \bar{P}_{u_c} \\ \bar{P}_{v_c} \\ \bar{P}_{w_c} \\ 1 \end{bmatrix} = \frac{1}{P_{Z_k}} \begin{bmatrix} -c & 0 & 0 & 0 \\ 0 & -c & 0 & 0 \\ 0 & 0 & -c & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} P_{X_k} \\ P_{Y_k} \\ P_{Z_k} \\ 1 \end{bmatrix} \quad (2.7)$$

Hierbei werden die Ortsvektoren zu den bezeichneten Punkten üblicherweise in homogenen Koordinaten angegeben. Auf die Einführung von homogenen Koordinaten wird an dieser Stelle nicht weiter eingegangen und auf die einschlägige Literatur verwiesen. Bei der Abbildung eines Punktes $P \in \mathbb{R}^3$ auf den zweidimensionalen Raum der Bildebene n geht die Tiefeninformation verloren. Bei der Abbildungsvorschrift kann also die Zeile, welche die Z -Achse beschreibt, vernachlässigt werden. Damit ergibt sich die Projektionsmatrix \mathbf{A} für die Abbildung mit einer Lochkamera nach Gleichung 2.8

$$\mathbf{A}_1 = \begin{bmatrix} -c & 0 & 0 \\ 0 & -c & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.8)$$

optischen Achse A parallel verschoben. Der Versatz der Punkte \bar{P} und P' beruht auf Abbildungsfehlern des Linsensystems der Kamera. Die Abbildungsfehler einer Linse werden in Abschnitt 2.4.1 erläutert. Der rotatorische Versatz der Bildebene S_c zum Koordinatensystem S_K der Kamera wird im Allgemeinen als klein angenommen und vernachlässigt. Zudem kann davon ausgegangen werden, dass der Abstand der Hauptebenen H_1 und H_2 im Verhältniss zu der Entfernung der Linse zur Bildebene auch klein ist. Somit können die Hauptebenen H_1 und H_2 als koplanar angesehen werden. Alle Abbildungsstrahlen treten dann kollinear durch das Projektionszentrum K_{Pz} der Linse. Für K_{Pz} gilt Gleichung 2.9.

$$K_{Pz} = K_1 = K_2 \quad (2.9)$$

K_{Pz} bildet den Ursprung O_{S_K} des Koordinatensystems S_K der Kamera. Die Z -Achse des Kamerakoordinatensystems wird entlang der optischen Achse in positiver Blickrichtung der Kamera definiert. Die X - und Y -Achse des Kamerakoordinatensystems liegen in der Bildebene und bilden mit der Z -Achse ein kartesisches Rechtssystem.

2.4.1 Die Abbildungsfehler einer Linse

Eine geometrisch ideale Bikonvexlinse hat ein Abbildungsverhalten, das dem der Lochkamera gleicht. Aufnahmen mit realen Linsen werden jedoch durch Abweichungen der Linsenform von ihrer idealen Form beeinflusst. Dies hat zur Folge, dass die Brennweite einer realen Linse vom durchleuchteten Flächenelement abhängig ist. Die Abbildung 2.5 zeigt den Strahlengang durch eine reale Linse.

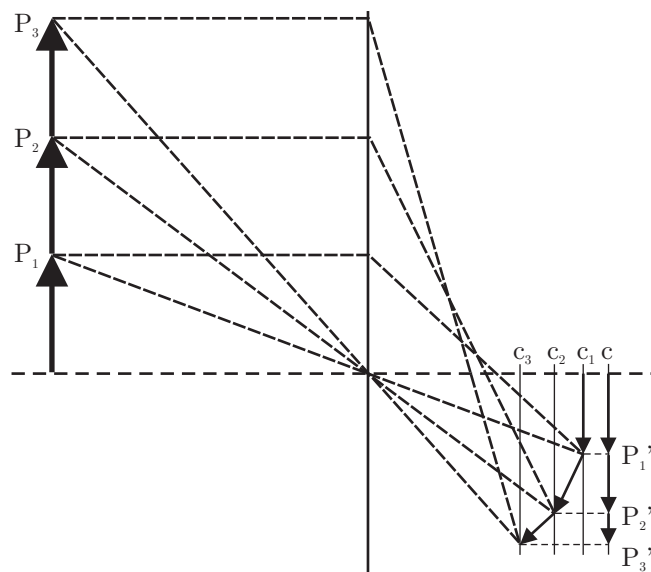


Bild 2.5: Strahlengang durch eine reale Linse

Bei einem optischen System werden die radial-symmetrische, die radial-asymmetrische und die tangentiale Verzeichnung unterschieden. Bei der radial-symmetrischen-Verzeichnung treten in Abhängigkeit der Linsenform zum Beispiel kissen- oder fassförmige Verzeichnungen auf. Die Abbildung 2.6 zeigt unterschiedliche Formen der Verzeichnung.

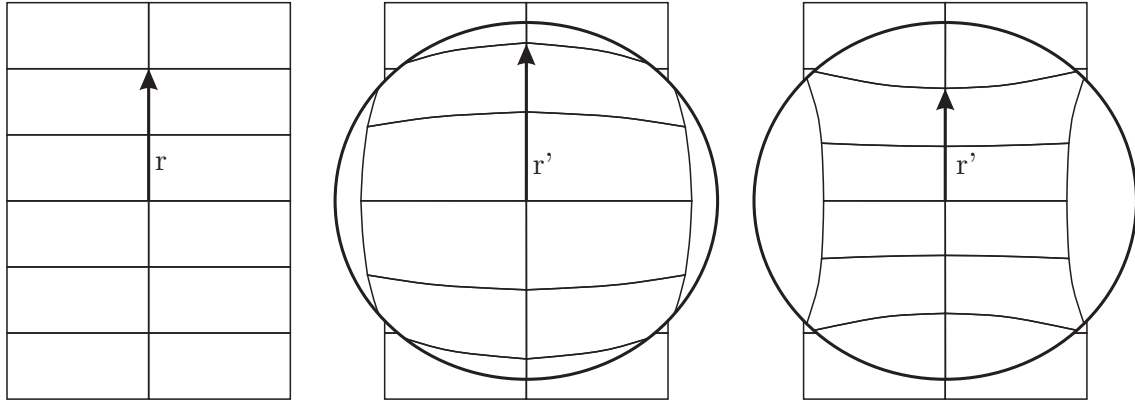


Bild 2.6: Radiale Linsen-Verzeichnung

2.4.2 Die radiale Verzeichnung

Die radiale Linsenverzeichnung hat zur Folge, dass ein Punkt auf dem Radius nach außen oder innen versetzt abgebildet wird. Sie lässt sich mathematisch als Polynom mit geradzahligem Exponenten modellieren [7]. In Abbildung 2.6 wird der Effekt der radialen Verzeichnung deutlich. Es wird vereinfacht davon ausgegangen, dass die Linse über einen Kreisring ihrer durchleuchteten Fläche eine konstante Fokusslänge besitzt. Das gemeinsame Zentrum dieser Kreisringe kann analog zur Lochkamera als optisches Zentrum angesehen werden. Hier wird der Ursprung des Kamerakoordinatensystems definiert. Alle Punkte, die auf einem Radius r um den Ursprung des Kamerakoordinatensystems S_k liegen, werden demnach mit der gleichen radialen Verzeichnung auf einen neuen Radius \tilde{r} abgebildet. Der verzeichnete Radius \tilde{r} ergibt sich nach Gleichung 2.10.

$$\tilde{r} = r + k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots \quad (2.10)$$

mit

$$r^2 = x^2 + y^2. \quad (2.11)$$

Die Abbildung 2.7 zeigt wie ein Bildpunkt radial verzeichnet abgebildet wird.

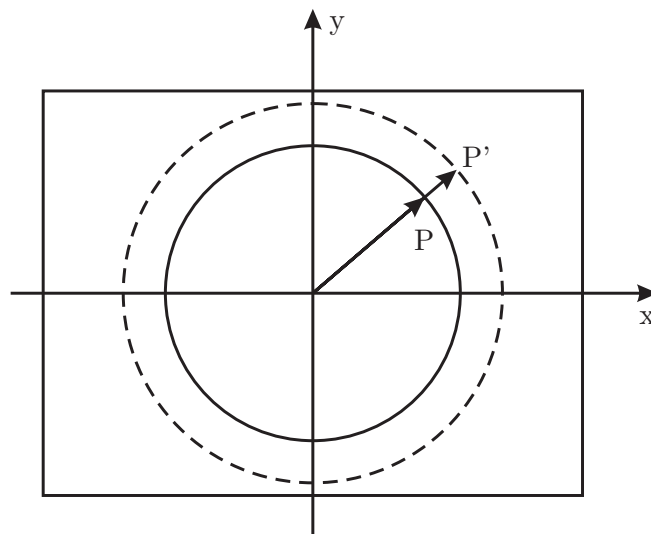


Bild 2.7: Prinzip der radialen Verzeichnung

Der Radius r zu einem Punkt $P|_{S_c}$ ergibt sich aus der Anwendung des Satzes von Pythagoras. Nach Tsai [7] erhält man für die Berechnung der radialen Verzeichnung bereits mit einem Polynom vierten Grades eine hinreichend gute Näherung. Die Abbildung 2.8 zeigt den Pixelversatz, der durch die radiale Verzeichnung auftritt.

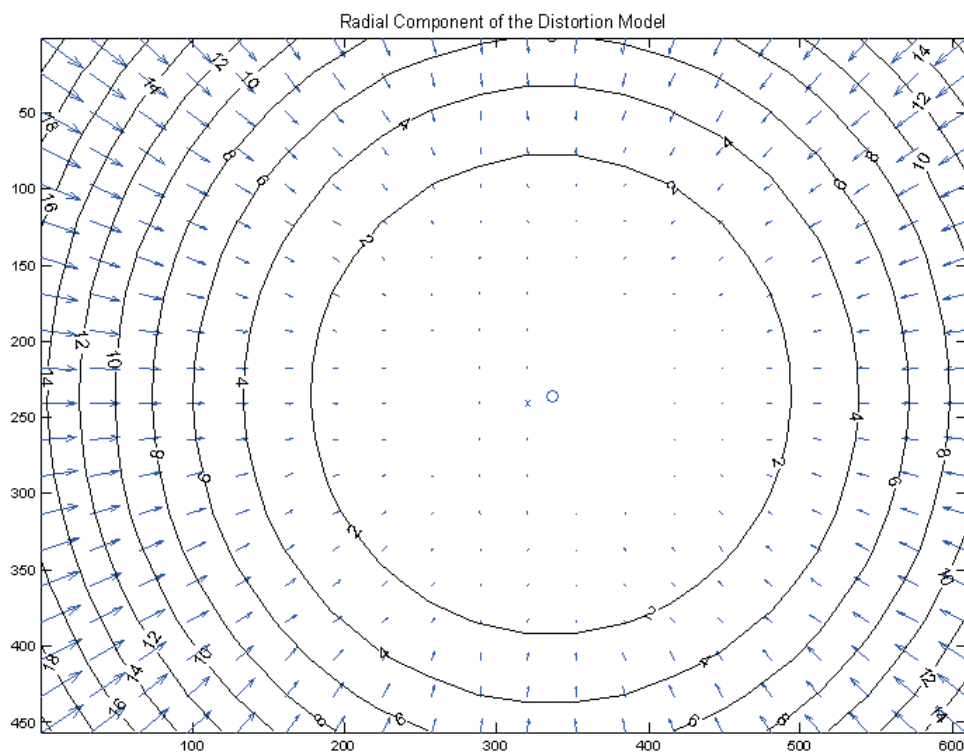


Bild 2.8: Radiale Verzeichnung

2.4.3 Die tangentiale Verzeichnung

Die tangentiale Verzeichnung hat zur Folge, dass der Punkt $P|_{S_c}$ entlang der Tangente an den Kreis K_P durch den Punkt $P|_{S_c}$ versetzt wird. Der Kreis K_P um den Ursprung des Kamerakoordinatensystems wird wiederum durch den abzubildenden Punkt $P|_{S_c}$ definiert. Dieser Versatz ist keine Funktion des Radius. Hierfür werden die Verzeichnungskonstanten p_1 und p_2 direkt mit dem betrachteten Bildpunkt $P|_{S_k}$ verknüpft. Die Abbildung 2.9 verdeutlicht wie ein Bildpunkt tangential verzeichnet abgebildet wird.

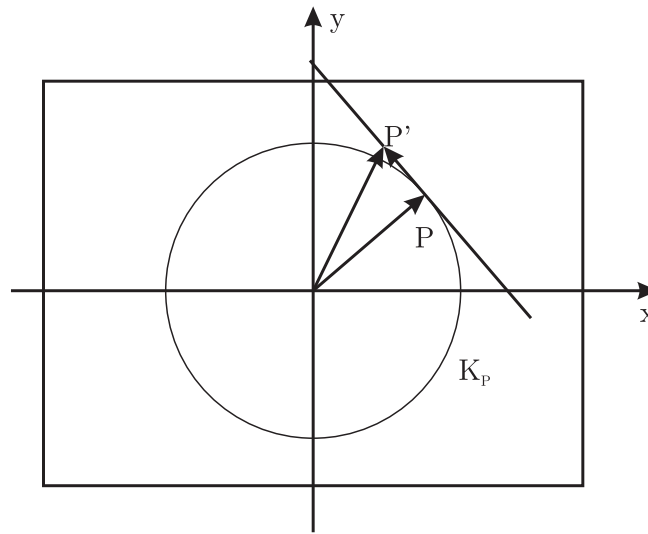


Bild 2.9: Prinzip der tangentialen Verzeichnung

Im Gegensatz zur radialen Verzerrung ist der Einfluss der tangentialen Verzerrung nach Zhang [22] sehr gering. Der Einfluss der tangentialen Verzerrung wird für die weiteren Betrachtungen vernachlässigt. Die tangentiale Verzeichnung kann durch die Gleichungen 2.12 und 2.13 beschrieben werden.

$$\tilde{x} = x + x(2p_1xy + p_2(r^2 + 2x^2)) \quad (2.12)$$

$$\tilde{y} = y + y(2p_2xy + p_1(r^2 + 2y^2)) \quad (2.13)$$

Es wird deutlich, dass im Gegensatz zur radialen Verzeichnung der Effekt der tangentialen Verzeichnung nicht rotationssymmetrisch zum optischen Zentrum der Linse ist.

Die Abbildung 2.10 zeigt den Pixelversatz, der durch die tangentiale Verzeichnung auftritt.

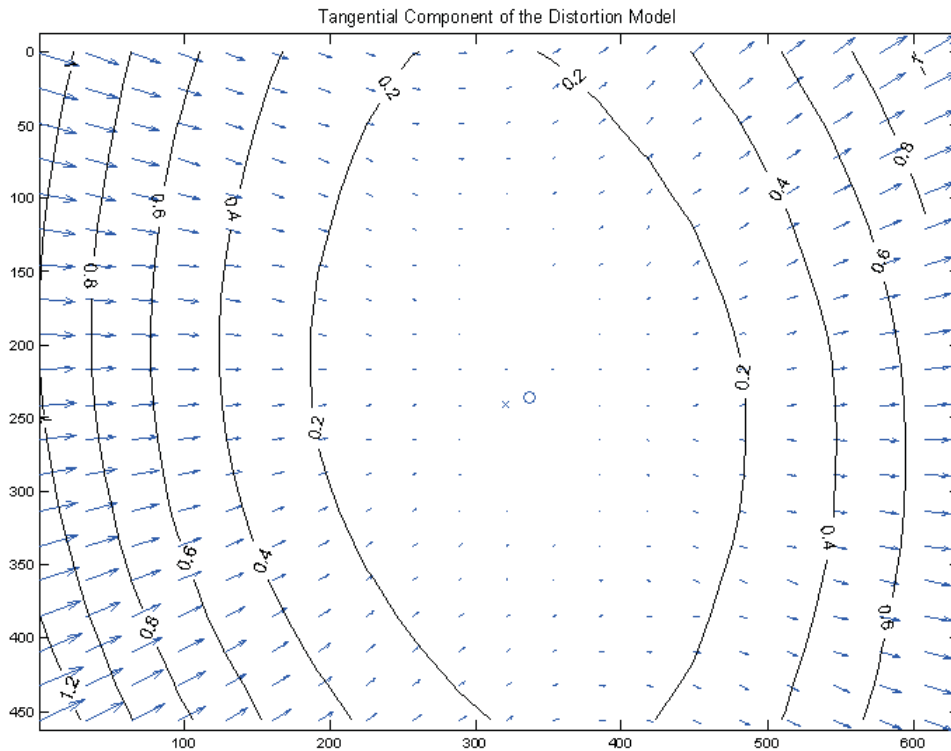


Bild 2.10: Tangentiale Verzeichnung

2.4.4 Die Abbildung auf einen realen Bildsensor

Wie schon in Abschnitt 2.4 erwähnt, sind die Hauptebenen der Linse sowie der Sensorchip, der die Bildebenen der Kamera darstellt, im Allgemeinen nicht planparallel. Zudem muss davon ausgegangen werden, dass der Ursprung des Sensorchips nicht auf der optischen Achse des Objektivs liegt. Es ist üblich, dass der Adressbereich der Kamerasensoren in der unteren linken Ecke des Sensors beginnt. Hier befindet sich also der informationstechnische Ursprung des Koordinatensystems des Sensors. Die Bildinformationen werden dann zeilenweise von links nach rechts und von unten nach oben im Speicher abgelegt. Diese Datendarstellung wird als *Row-Major-Format* bezeichnet. Die spaltenweise Darstellung der Daten wird analog als *Column-Major-Format* bezeichnet und entspricht der Transponierung des Row-Major-Formats. Die Zeilen des Sensors verlaufen parallel zur X -Achse, die Spalten parallel zur Y -Achse des Sensorkoordinatensystems. Hierbei zeigt sowohl die X - als auch die Y -Achse in Richtung aufsteigender Adresswerte. Der sich daraus ergebende Normalenvektor auf die Bildebene $\vec{n}_{S_c} = \vec{e}_{u,S_c} \times \vec{e}_{v,S_c}$ zeigt in positive Aufnahme­richtung. Somit ist die Konvention nach einem kartesischen Rechtssystem erfüllt. An dieser Stelle sei erwähnt, dass nicht alle derzeit verfügbaren Kameras dieser Koordinatensystemkonvention entsprechen.

Die Transformation eines Punktes $P|_{S_c}$ aus dem in Abschnitt 2.3 beschriebenen Koordinatensystem der idealen Bildebene in das Koordinatensystem eines Sensorchips kann wiederum durch eine Translation und eine Rotation um die drei Raumwinkel erreicht werden. Die Abbildung des Punktes $P|_{S_c}$ von der idealen Bildebene auf den Punkt $\bar{P}_{S_{sc}}$ des Sensorchips ergibt sich gemäß der Gleichungen 2.1 und 2.7 nach Gleichung 2.14.

$$\bar{P}|_{S_{sc}} = [\mathbf{R}_{sc} \vec{t}_{sc}] P|_{S_c} \quad (2.14)$$

Hierbei ergibt sich die Rotationsmatrix \mathbf{R}_{sc} nach Gleichung 2.2. Die Winkel $\theta_x, \theta_y, \theta_z$ bemessen die Drehung des Sensorschips bezüglich der idealen Bildebene um seine X -, Y - und Z -Achse. Der Vektor \vec{t}_{sc} beschreibt den translativen Versatz zwischen dem Ursprung des Kamerakoordinatensystems und dem Ursprung des Sensorkoordinatensystems. Der Vektor \vec{t}_{sc} ergibt sich nach Gleichung 2.15.

$$\vec{t}_{sc} = \vec{O}_{sc}|_{S_0} - \vec{O}_c|_{S_0} \quad (2.15)$$

Die Vektoren $\vec{O}_{sc}|_{S_0}$ und $\vec{O}_c|_{S_0}$ bezeichnen den Ursprung der idealen Bildebene S_k und des Koordinatensystems des Sensors S_{sc} im Weltkoordinatensystem S_0 . Der Einfluss des rotatorischen Versatzes des Sensorchips bezüglich seiner Ideallage wird im Allgemeinen als klein eingeschätzt und vernachlässigt. So kann die Rotationsmatrix \mathbf{R}_{sc} als Einheitsmatrix angenommen werden.

2.4.5 Die Abbildung auf einen diskretisierenden Sensor

Beim Bildeinzug mit einer digitalen Kamera wird das auf der Bildebene projizierte Bild über der Länge und Breite des Sensors diskretisiert. Darüber hinaus wird auch die Farbdarstellung diskretisiert. Der hierfür relevante Sachverhalt der Diskretisierung der Farben wird in Kapitel 3.3.3 erläutert.

Im Allgemeinen besitzen die sensitiven Elemente des Sensorchips keine quadratische Oberfläche. Somit liefert der Sensorchip unterschiedliche Auflösungen entlang seiner X - und Y -Achse auf. Die sensitiven Elemente vieler CCD-Sensorchips weisen eine Scherung s auf. Der Einfluss der Scherung wird als gering eingeschätzt und vernachlässigt. Zudem sind die Pixel des Sensors mit einem Mindestabstand zueinander angeordnet. Vereinfachend wird angenommen, dass die Pixel des Sensorchips bündig aneinandergereihte Flächenelemente sind. Diese Vereinfachung ist in sofern zulässig, als der Sensor in den Bereichen zwischen den Pixeln keine Information über die lokale Belichtung liefert. Die Abbildung 2.11 zeigt den Übergang vom

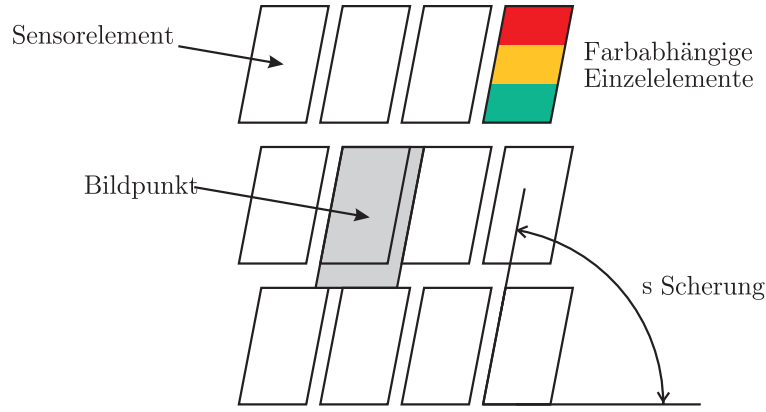


Bild 2.11: Modell des Sensorchips

realen Sensorchip zum allgemein verwendeten Modell eines CCD-Sensorchips. Es ist üblich, dass Längenangaben bei der Bemessung von Kamerasystemen zunächst in Längeneinheiten, die der Kantenlänge eines Sensorelementes entsprechen, angegeben werden. Bezeichnet man die Anzahl der Pixel pro Längeneinheit parallel zur u -Achse des Sensorkoordinatensystems mit k_u und entlang der v -Achse mit k_v , so ergeben sich die Koordinaten u_{sc} und v_{sc} des Pixels $\bar{P}|_{sc}$ zu einem $P|_c$ der Bildebene nach Gleichung 2.16

$$k_u u_c + u_{sc,0} = u_{sc} \quad (2.16)$$

$$k_v v_c + v_{sc,0} = v_{sc}$$

$$(u_{sc,0} \leq u_{sc,m} \leq u_{sc,max}) \in \mathbb{Z}$$

$$(v_{sc,0} \leq v_{sc} \leq v_{sc,max}) \in \mathbb{Z}$$

Hierbei bezeichnen die Pixel $u_{sc,0}$ und $v_{sc,0}$ den Ursprung des Sensors und berechnen sich nach Gleichung 2.17

$$u_{sc,0} = -k_u u_{c,min}, \quad v_{sc,0} = -k_v v_{c,min} \quad (2.17)$$

$$u_{sc,max} = -k_u u_{c,max}, \quad v_{sc,max} = -k_v v_{c,max} \quad (2.18)$$

Eine kameraspezifische Längeneinheit berechnet sich dann nach Gleichung 2.19.

$$l_u = \frac{k_u (u_{c,max} - u_{c,min})}{(u_{sc,max} - u_{sc,0})} \quad (2.19)$$

$$l_v = \frac{k_v (v_{c,max} - v_{c,min})}{(v_{sc,max} - v_{sc,0})} \quad (2.20)$$

Die Koordinaten des abgebildeten Punktes $\vec{P}|_{s_{sc}}$ auf eine planar um den Vektor $\vec{t}_{sc} = [u_{sc,0} v_{sc,0}]^T$ verschobene Bildebene mit unterschiedlich skalierten Achsen berechnen sich somit nach Gleichung 2.21

$$\vec{P}|_{s_{sc}} = \begin{pmatrix} k_u \\ k_v \end{pmatrix} \vec{P}_c^* + \vec{t}_{sc} \quad (2.21)$$

Die Projektionsmatrix \mathbf{A}_2 für die Transformation eines Punktes \vec{P}^* in das diskrete Koordinatensystem S_{sc} des Sensorchips lautet dann:

$$\mathbf{A}_2 = \begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.22)$$

Die Abbildung durch eine Lochkamera auf einen Sensorchip wird durch die Projektionsmatrix \mathbf{A} , die sich durch die Multiplikation der Matrizen \mathbf{A}_1 und \mathbf{A}_2 ergibt, beschrieben.

$$\mathbf{A} = \mathbf{A}_2 \mathbf{A}_1$$

$$\mathbf{A} = \begin{bmatrix} -ck_u & 0 & u_{sc,0} \\ 0 & -ck_v & v_{sc,0} \\ 0 & 0 & 1 \end{bmatrix} \quad (2.23)$$

Die allgemeine Abbildung $\vec{P}|_{s_{sc}}$ des Punktes $P|_{s_0}$ durch eine Lochkamera auf einen Sensorchip ergibt sich dann unter Berücksichtigung der Gleichung 2.14 nach:

$$P|_{s_{sc}} = \mathbf{A}[\mathbf{R}_K \vec{t}_K] P|_{s_0} \quad (2.24)$$

$$A_{sc} = \mathbf{A}[\mathbf{R}_K \vec{t}_K] \quad (2.25)$$

Die Abbildungsmatrix \mathbf{A}_{sc} beschreibt die Abbildung von Objektpunkten auf dem Kamerasensor unter Berücksichtigung des Versatzes des Kamerakoordinatensystems zum Weltkoordinatensystem und des Abbildungsverhaltens der Kamera.

2.4.6 Farbdarstellung

In diesem Abschnitt werden die grundlegenden Begriffe der Farbdarstellung durch Digitalkameras erläutert. Die informationstechnische Grundlage des *RGB-Formates* bildet das gleichnamige Farbmodell, das zur Darstellung von Farben dient. Es verwendet nur die drei Grundfarben Rot, Grün und Blau. Durch additive Mischung der Grundfarben lassen sich alle anderen Farbtöne des sichtbaren Lichtes erzeugen. Schwarz und Weiß entstehen dabei

durch vollständige Abwesenheit bzw. vollständige Überlagerung der Grundfarben. Die Helligkeitsinformation ist in den Farbtönen enthalten. Üblicherweise werden die Intensitäten bei der Verwendung des RGB-Farbmodells auf 256 Werte abgebildet. Tabelle 2.1 zeigt einige Beispiele zur Farbmischung. Die zu mischenden Farben stehen links und die Grundfarben oben.

Farbton	Rot	Grün	Blau
Rot	255	0	0
Grün	0	255	0
Blau	0	0	255
Schwarz	0	0	0
Weiß	255	255	255

Tabelle 2.1: Farbdarstellung mit dem RGB-Farbmodell

2.4.7 Sättigung und Luminanz

Der Sättigungsgrad eines Pixels gibt den Grad der Ausnutzung des Wertebereichs der Intensität eines Pixels im betrachteten Merkmal an. Die Luminanz ist die Helligkeit des dargestellten Farbtons unabhängig vom Farbton. Sie ist somit direkt abhängig von der Sättigung der beteiligten Komponenten. Es sei angemerkt, dass das Verhältnis der Farbanteile zueinander hierdurch nicht beeinflusst wird.

2.5 Stereo-optische Systeme

Nachdem in den vorangegangenen Kapiteln auf die mathematischen Zusammenhänge bei einer Kamera eingegangen wurde, findet im Abschnitt 2.6 eine ausführliche Betrachtung der geometrischen Beziehung zwischen zwei unterschiedlichen Kameras statt. Solche Zwei-Kamera-Anordnungen werden Stereosystem oder Epipolarsystem genannt. Epipolarsysteme werden in achsparallele und konvergente Systeme unterschieden.

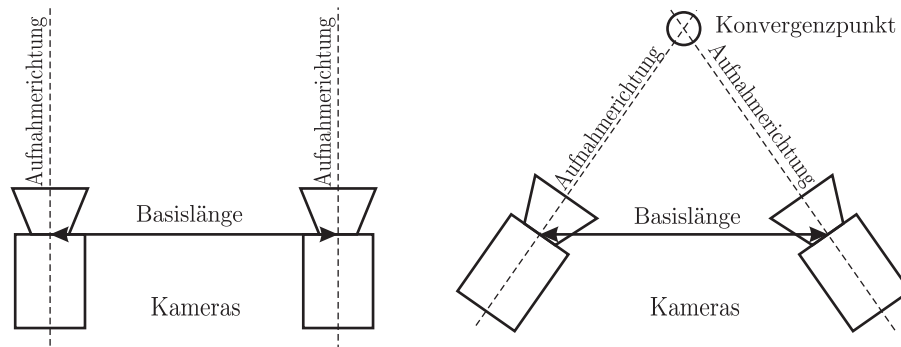


Bild 2.12: Achsparalleles Stereosystem (links), konvergente Anordnung (rechts)

Die Abbildung 2.12 zeigt sowohl ein achsparalleles als auch ein konvergentes Stereosystem. Das achsparallele System ist durch eine parallele Ausrichtung der optischen Achsen beider Kameras gekennzeichnet, wohingegen die konvergente Anordnung die optischen Achsen auf einen Konvergenzpunkt ausrichtet. Die mathematischen Zusammenhänge des achsparallelen Stereosystems sind einfacher zu beschreiben, weisen aber erhebliche Einschränkungen auf, da die Kameras nicht auf ein bestimmtes Objekt ausgerichtet sind. Unter Umständen kann die Szene nicht vollständig von beiden Kameras erfasst werden. In der Praxis wird meist vom konvergenten Aufbau ausgegangen, da es nicht möglich ist, ohne großen Aufwand ein achsparalleles System aufzubauen.

2.6 Die Epipolargeometrie

Im allgemeinen Fall sind die Kameras eines Stereosystems zueinander verschoben und gedreht. Ein Punkt $P|_{S_0}$ im Weltkoordinatensystem kann mit Gleichung 2.1 in die Koordinatensysteme der beteiligten Kameras transformiert werden. Die Transformation eines Punktes

aus dem Koordinatensystem der linken Kamera in das Koordinatensystem der rechten Kamera beschreibt Gleichung 2.26.

$$P|_{S_{kR}} = [\mathbf{R}_R \vec{t}_R] P|_{S_0} \quad (2.26)$$

$$P|_{S_{kL}} = [\mathbf{R}_L \vec{t}_L] P|_{S_0} \quad (2.27)$$

$$P|_{S_{kR}} = [\mathbf{R}_R \vec{t}_R][\mathbf{R}_L \vec{t}_L]^{-1} P|_{S_{kL}} \quad (2.28)$$

Das Problem wird einfacher, wenn das Weltkoordinatensystem für das Stereosystem gleich dem Koordinatensystem einer der beiden Kameras gewählt wird. Dieser Ansatz ist sinnvoll, da im Allgemeinen eine der beiden Kameras als Referenzsystem oder Master-Kamera bestimmt werden muss. Für alle weiteren Betrachtungen wird die linke Kamera als Referenzkamera bestimmt. Definiert das Koordinatensystem der linken Kamera das Weltkoordinatensystem, so ergibt sich die Transformation des Punktes $P|_{S_{kL}}$ aus dem Koordinatensystem der linken Kamera in das Koordinatensystem der rechten Kamera nach Gleichung 2.29

$$P|_{S_{kR}} = [\mathbf{R}_R \vec{t}_R] P|_{S_{kL}} \quad (2.29)$$

Die Abbildung 2.13 zeigt die Epipolargeometrie eines Stereo-Kamera-Systems für einen Punkt P .

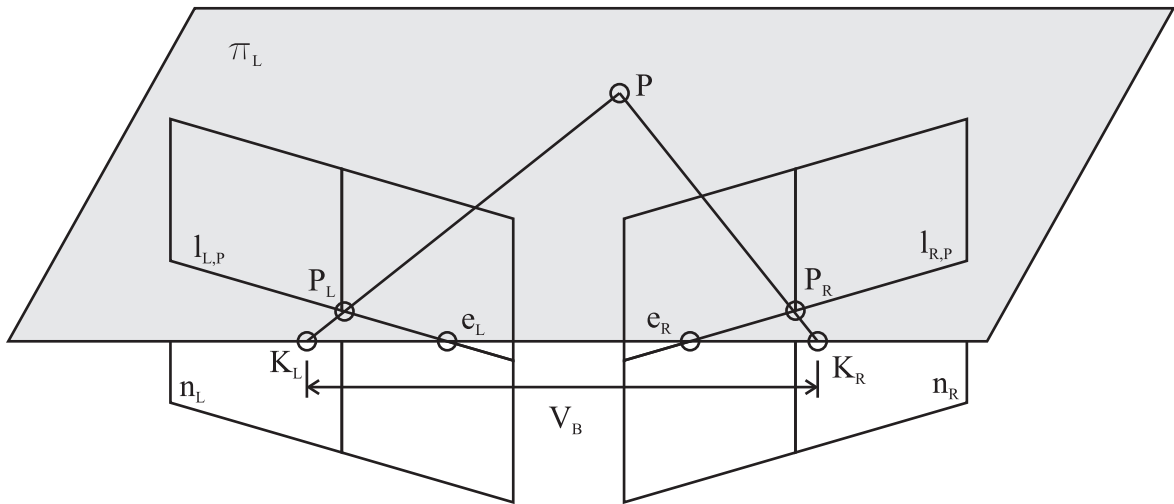


Bild 2.13: Allgemeine Epipolargeometrie

Die Bildpunkte P_L und P_R werden vom Punkt P auf die linke und rechte Bildebene n_L und n_R abgebildet. Diese Punkte werden als korrespondierendes Punktpaar bezeichnet. Der Versatz eines abgebildeten Punktes P_L im linken Bild bezüglich des korrespondierenden Punktes P_R

im rechten Bild wird Disparität D_P genannt. Sie berechnet sich nach Gleichung 2.30

$$\overrightarrow{D_P} = \overrightarrow{P_L} - \overrightarrow{P_R} \quad (2.30)$$

$\overrightarrow{t_R}$ bezeichnet den Verbindungsvektor der optischen Zentren der linken und rechten Kamera. $\mathbf{R_R}$ ist die Rotationsmatrix, die den rotatorischen Versatz der rechten bezüglich der linken Kamera beschreibt. Die Verbindungsgerade zwischen den beiden optischen Zentren wird Basislinie V_B genannt. Sind die Bildebenen n_L und n_R planparallel verschoben, so ist die Basislinie Element beider Bildebenen. Liegt ein rotatorischer Versatz der Bildebenen vor, so durchdringt die Basislinie V_B die Bildebenen an den Punkten e_L und e_R . Die Punkte e_L und e_R heißen *Epipole*. Ihre Lage in den Bildebenen ist nur durch die Anordnung der Kameras zueinander bestimmt. Die Epipole sind zudem die Projektion der optischen Zentren einer Kamera auf die Bildebene der jeweils anderen Kamera. Der Objektpunkt $P \in \mathbb{R}^3$ und die beiden Brennpunkte der Kameras spannen die Epipolarebene π_p für diesen Punkt auf (vgl. Abbildung 2.13). Die Abbildungen P_L und P_R sowie deren Abbildungsstrahlen sind ebenfalls Element der Epipolarebene π_p . Die Schnittgeraden der Epipolarebene π_p mit den Bildebenen n_L und n_R sind die *Epipolarlinien* $l_{L,P}$ und $l_{R,P}$ zu dem Punkt P . Die Epipolarlinie $l_{R,P}$ im rechten Bild zu einem Punkt P_L im linken Bild ist zudem die Projektion des zu diesem Punkt gehörenden Abbildungsstrahls und umgekehrt. Somit definiert die Epipolarlinie $l_{R,P}$ die Menge der Bildpunkte aus n_r , deren Element der korrespondierende Punkt P_R zu einem Punkt P_L aus n_L ist. Für den korrespondierenden Punkt P_R zu einem Punkt P_L gilt die Bedingung 2.31

$$P_R \in l_{RP} \quad (2.31)$$

2.7 Die Homographiematrix H

Als Homographie wird die Abbildung von Punkten auf gleich geartete Räume bezeichnet. In dieser Arbeit wird im Rahmen der Kalibrierung des Kamerasystems die Abbildung von Punkten eines planaren Kalibrierungsmusters auf die Bildebenen der Kameras betrachtet. Darüber hinaus wird die Abbildung eines Objektes in mehreren optischen Systemen sowie in mehreren Bildern einer Sequenz behandelt. Bezeichnet man die korrespondierenden Punkte der linken und rechten Kamera mit P_L und P_R so gilt Gleichung 2.32.

$$P_R = \mathbf{H_{LR}} P_L \quad (2.32)$$

\mathbf{H}_{ik} bezeichnet die Homographiematrix. Unter Berücksichtigung der Abbildungsmatrizen \mathbf{A}_i der beteiligten Kameras ergibt sich die Homographiematrix \mathbf{H}_{LR} für die korrespondierenden Punkte P_L und P_R eines Stereovisionssystems zu einem Objektpunkt P im allgemeinen Weltkoordinatensystem aus Gleichung 2.33.

$$\mathbf{H}_{LR} = \mathbf{A}_R[\mathbf{R}_R \vec{t}_r](\mathbf{A}_L[\mathbf{R}_L \vec{t}_L])^{-1} \quad (2.33)$$

Definiert man nun das Weltkoordinatensystem wieder im Koordinatensystem der linken Kamera, vereinfacht sich der Ausdruck der Homographie analog zu Gleichung 2.29 zu Gleichung 2.34.

$$\mathbf{H}_{LR} = \mathbf{A}_R[\mathbf{R}_R \vec{t}_r]\mathbf{A}_L^{-1} \quad (2.34)$$

Betrachtet man nun zusätzlich die Abbildung einer Ebene in den zwei Kameras, so können über die Homographie der Objektebenen zu den beteiligten Bildebenen die Abbildungsmatrizen \mathbf{A}_i sowie die Rotationsmatrix und der Translationsvektor zwischen den Bildebenen bestimmt werden. Somit stellt die Homographie eine Schlüsselinformation zur Kalibrierung eines Kamerasystems dar. Im Weiteren kann die Homographie auch den Versatz zweier korrespondierender Bildpunkte aus den Bildern einer Sequenz beschreiben. In dieser Verwendung bietet die Homographie die Möglichkeit, die Bilder einer Sequenz zu einem Gesamtbild zusammenzuführen.

2.8 Epipolargeometrische Beziehungen in Kamerakoordinaten

Bei der Anwendung der Stereovision zur Entfernungsbestimmung ist jedoch die Position des Objektpunktes P nicht bekannt. Bekannt sind in diesem Fall nur die Abbildungen P_L und P_R des Objektpunktes in den Koordinatensystemen der jeweiligen Kameras. Es ist also sinnvoll, die bis hierhin aufgeführten Beziehungen zusätzlich in Kamerakoordinaten zu betrachten.

2.8.1 Die Essentialmatrix E

Die Kamerakoordinatensysteme stehen über eine euklidische Transformation in Beziehung. Aus dem Lochkameramodell ergeben sich die Abbildungen P_L und P_R eines Objektpunktes

P in Sensorkoordinaten nach Gleichung 2.35.

$$P_L = \frac{P|_{S_{SC,L}}}{Z_L} \quad (2.35)$$

$$P_R = \frac{P|_{S_{SC,R}}}{Z_r} \quad (2.36)$$

Hierbei bezeichnet Z_i die bisher nicht bekannte Entfernung des Objektpunktes P einer Kamera i . Z wird im Weiteren als Strukturparameter bezeichnet. Durch die Kombination von Gleichung 2.29 und Gleichung 2.35 erhält man die Beziehung zwischen den Abbildungen P_L und P_R des Objektpunktes P .

$$P_R = \frac{1}{Z_R}(Z_L \mathbf{R} P_L + \vec{t}) \quad (2.37)$$

Erweitert man die Gleichung 2.37 um das beidseitige Kreuzprodukt mit dem Vektor \vec{t} und erweitert man das Skalarprodukt mit P_R , so erhält man nach Eliminierung der Strukturparameter die Gleichung 2.38.

$$P_R \cdot (\vec{t} \times R_R) = P_R^T \vec{t} \times (\mathbf{R} P_L) = 0 \quad (2.38)$$

Definiert man nun eine Zuordnung eines dreidimensionalen Vektors zu einer antisymmetrischen Matrix entsprechend Gleichung 2.39, so kann das Kreuzprodukt zweier Vektoren durch das Matrix-Vektor-Produkt einer 3×3 -Matrix durch einen dreidimensionalen Vektor ausgedrückt werden.

$$\begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}_{\times} = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} \quad (2.39)$$

mit

$$\mathbf{t}_{\times} = -\mathbf{t}_{\times}^T$$

Mit dieser Schreibweise erhält man schließlich die *Epipolargleichung* (2.40) für die mathematische Beziehung zwischen den Abbildungen eines Objektpunktes in zwei Kameras.

$$P_R^T \mathbf{E} P_L = 0 \quad (2.40)$$

$$\mathbf{t}_{\times} \mathbf{R} = \mathbf{E} \quad (2.41)$$

Diese Gleichung ist erfüllt, wenn die Abbildungen in den beteiligten Kameras korrespon-

dierende Abbildungen eines Objektpunktes sind. Es sei angemerkt, dass der Umkehrschluss nicht gilt. Abbildungen in beiden Kameras, welche die Epipolarbedingung erfüllen, sind demnach nicht zwangsläufig korrespondierende Punkte. Dies zeigt sich daran, dass alle Punkte auf einem optischen Strahl zu einer Abbildung in einer Ansicht und zu unterschiedlichen Abbildungen in der zweiten Ansicht führen. Die Bedeutung der Epipolargleichung wird nach der folgenden Herleitung der Epipolarlinien noch deutlicher. Die Essentialmatrix E beschreibt die euklidische Transformation von einer Kamera zu einer anderen Kamera des Epipolarsystems.

2.8.2 Die Epipolarlinie l

Eine Gerade in einer Ebene kann durch den Koeffizientenvektor in Parameterdarstellung beschrieben werden. Die Gleichung 2.42 zeigt eine allgemeine Geradengleichung in Parameterdarstellung sowie den zugehörigen Koeffizientenvektor \vec{g} .

$$g : 0 = hx + iy + j \quad (2.42)$$

mit

$$\vec{g} = \begin{pmatrix} h \\ i \\ j \end{pmatrix}$$

Für die Punkte P_i^T , die Element einer Linie g sind, muss gemäß der projektiven Geometrie [Anhang A.6] Gleichung 2.43 gelten.

$$0 = P_i^T \times g \quad (2.43)$$

$$P_i = \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix}.$$

Über das Kreuzprodukt zweier Punkte P_1 und P_2 ist eindeutig eine Gerade definiert.

$$\vec{g} = P_1 \times P_2 \quad (2.44)$$

Mit den Kameraparametern lässt sich jeder Objektpunkt über den zugehörigen Bildpunkt in homogenen Koordinaten und einem Skalierungsfaktor λ , welcher der Entfernung zum

optischen Zentrum entspricht, beschreiben.

$$\begin{aligned}\lambda_i 1P_i &= P|_{SC,i} \\ \lambda_i &\in [0, \infty]\end{aligned}\tag{2.45}$$

Mit dem Ursprung des Weltkoordinatensystems im Koordinatensystem der linken Kamera, ergibt sich aus den Gleichungen 2.29 und 2.45 die Gleichung 2.46.

$$\lambda_R P_R = \lambda_L [\mathbf{R}_R \vec{t}_r] P_L \tag{2.46}$$

Nun kann die Epipolarlinie in der rechten Kamera über zwei ausgezeichnete Punkte berechnet werden. Dies ist zum einen der Epipol in Abbildung 2.13, der die Projektion des optischen Zentrums der linken Kamera in die Bildebene n_R der rechten Kamera darstellt. Er entspricht damit einer Skalierung mit $\lambda_1 = 0$. Man erhält den Epipol aus Gleichung 2.47, der bis auf einen Faktor dem Vektor \vec{t} entspricht.

$$e_2 = \vec{t} \tag{2.47}$$

Der zweite Punkt ist die Abbildung eines Punktes im Unendlichen in der rechten Kamera. Dies entspricht einer Skalierung von P_R mit $\lambda_1 = \infty$. Damit kann der Vektor \vec{t} vernachlässigt werden und man erhält die Projektion des Punktes $P_{R\infty}$.

$$P_{R\infty} = R P_L \tag{2.48}$$

Aus diesen beiden Abbildungen kann nun die Epipolarlinie in Ansicht 2 berechnet werden.

$$\vec{l}_R = e_R \times P_{R\infty} = \vec{t} \times R P_L = E P_L \tag{2.49}$$

Die korrespondierende Epipolarlinie in Abbildung 2.13 ist nach Gleichung 2.50 gegeben.

$$\vec{l}_L = E^T P_L \tag{2.50}$$

Analog gilt die hier angeführte Berechnung der Epipolarlinie für Wahl eines Punktes P_R aus der Bildebenen der rechten Kamera.

3 Der Bildeinzug

In diesem Kapitel werden die Details sowie die Anforderungen des Prozesses an die Bilderfassung und die Bildvorverarbeitung erläutert. Es wird zunächst der Teilprozess des Bildeinzugs beschrieben. Dann wird auf die Problematik des primären Bildeinzugs durch mehrere Kameras mit Hinsicht auf das zu messende Objekt eingegangen. Im Weiteren ein Verfahren vorgestellt, welches es erlaubt, bei natürlich variierender Belichtung Bilder mit annähernd gleich bleibender Qualität zu erfassen. Als Voraussetzung dafür wird eine Funktion entwickelt, die es ermöglicht, die Qualität von Bildern quantitativ zu bewerten. Neben der Steuerung der Hardware werden die verwendeten Methoden zur Optimierung der Bildqualität hinsichtlich ihrer Eignung zur stereoskopischen Auswertung beschrieben.

3.1 Der Prozess des Bildeinzug

Im Funktionsblock Bildeinzug werden die Bilddaten über die angeschlossenen Kameras eingezogen. Dazu werden die Kameras so angesteuert, dass die Bilder zeitgleich digitalisiert werden. Hierfür wird die Belichtungszeit t_{exp} und die Sensorverstärkung V_{ccd} , welche anhand der Bilder des vorhergehenden Berechnungszyklus $n - 1$ bestimmt wurden, verwendet. Nachdem die Bilder von der Treibersoftware im Speicher zur Verfügung gestellt wurden, wird die Qualität Q_{ist} der eingezogenen Bilder hinsichtlich des relevanten Merkmals ermittelt. Der Abschnitt 3.3.1 erläutert die Details der Bildqualität sowie deren Bemessung. Das Ergebnis der Ermittlung der Bildqualität ist der skalare Wert Q_{ist} . Die Regelabweichung E_Q wird einem Regelkreis zur Steuerung der Bildqualität zugeführt. Die Abbildung 3.1 zeigt das Flussdiagramm sowie den Signalfluss des Funktionsblockes für den Bildeinzug und des Belichtungsregelkreises. Die Ausgabe des Reglers wird einem Stellalgorithmus zur Berechnung der Belichtungszeit t_{exp} und der Sensorchip-Verstärkung V_{ccd} zugeführt. Die Details dieser Funktion sind auch in Abschnitt 3 angeführt. Die ermittelten Werte für die Stellgrößen t_{exp} und V_{ccd} werden im nächsten Prozessdurchlauf an die Kameras übertragen.

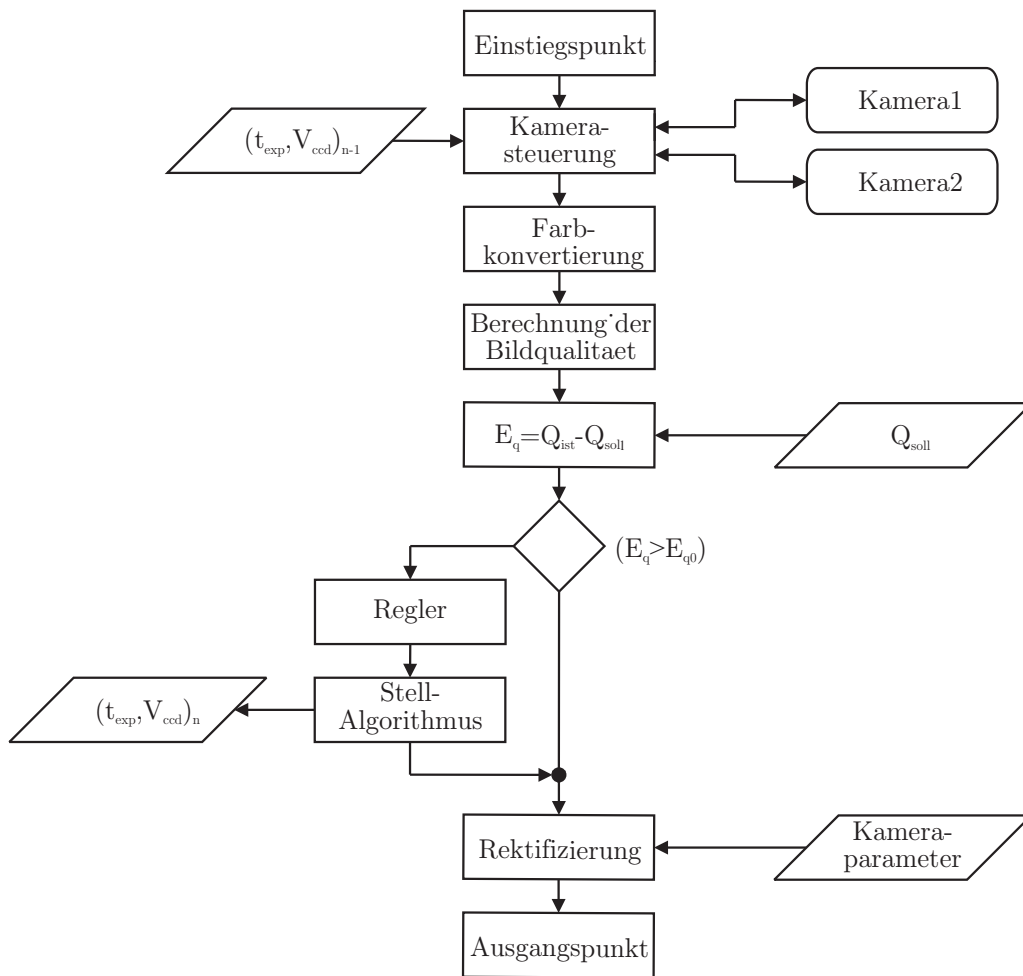


Bild 3.1: Flussdiagramm des Funktionsblockes Bildeinzug

Dies geschieht bevor die nächsten Bilder eingezogen werden. Bei der hier verwendeten Kameragerätetechnik ist die Neukonfiguration der Kameras mit einer Wartezeit von ca. $t_{lat} = 0,1s$ verbunden. Dies betrifft auch die Änderung der Belichtungszeit und der Sensorchip-Verstärkung. Da die Gesamtfunktion keine multitaskingbasierte Abarbeitung der Aufgaben vorsieht, ist die Änderung der Kameraparameter mit unnötiger Stillstandzeit des Prozesses verbunden. Diesem Effekt wurde Rechnung getragen, indem die Berechnung der Stellgröße sowie die Änderung der Kameraeinstellungen nur bei einer Überschreitung der Mindestregelabweichung $E_q > E_{q0}$ ausgelöst wird. Eine asynchrone, parallele Abarbeitung des Bildeinzuges und der weiteren Bildverarbeitung ist eine weitere Möglichkeit, um Stillstandszeiten des Prozesses zu reduzieren. Dieses wurde bisher aber nicht implementiert. Wird die Berechnung der Stellgrößen für die Kameras nicht angestoßen, so werden lediglich die Bilddaten durch die Kameras eingezogen und der weiteren Datenverarbeitung zur Verfügung gestellt.

3.2 Die Anforderungen an den Bildeinzug

Das Messsystem misst den Füllstand optisch. Hierbei wird die Oberfläche des Ladegutes mit Hilfe einer Stereovisionsanalyse vermessen und numerisch beschrieben. Im Rahmen der Stereovisionsanalyse wird die Disparität zweier korrespondierender Bildpunkte ausgewertet. Im folgenden Abschnitt 3.2.1 wird die Auswahl einer geeigneten Kameragerätetechnik diskutiert.

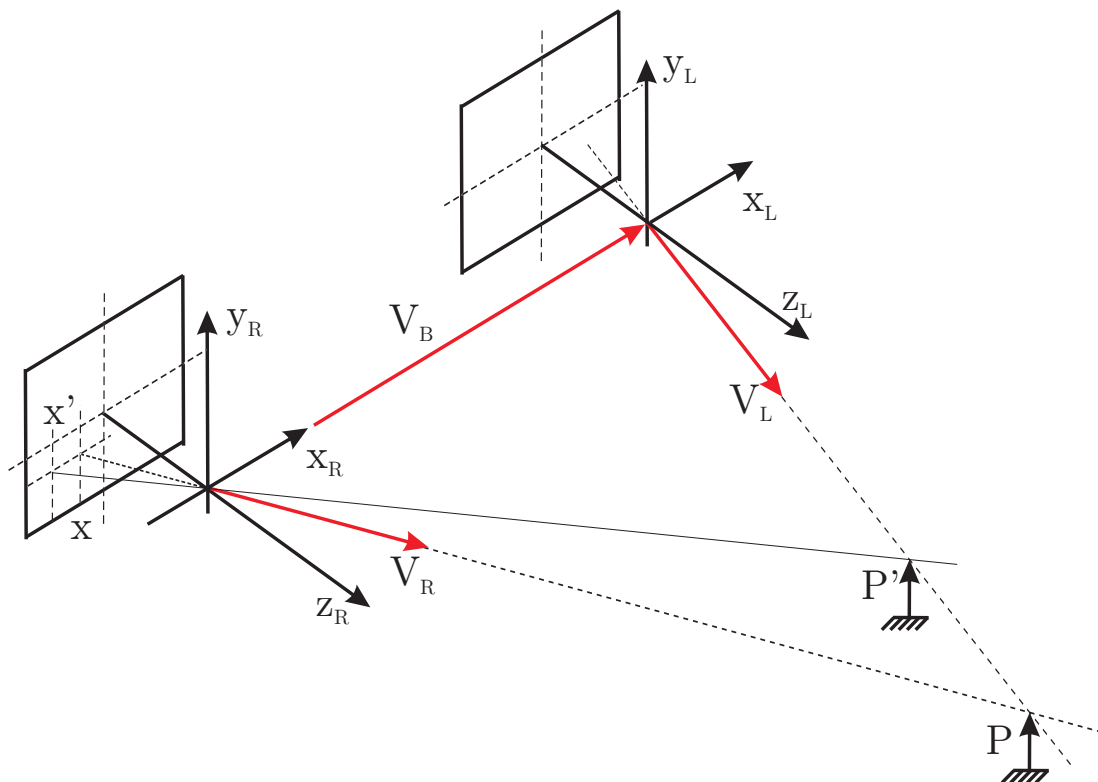


Bild 3.2: Abbildung in zwei optischen Systemen

Die Paare korrespondierender Punkte werden über einen Vergleich der jeweiligen Bildausschnitte in der Umgebung der betrachteten Punkte ermittelt. Dies setzt voraus, dass die eingezogenen Bilder einen hinreichend großen Informationsgehalt enthalten hinsichtlich ihrer Eignung, in der Bildebene eindeutig lokalisiert werden zu können. Auf die hier relevante Information wird in den Kapiteln 3.3.1 und 5.2 eingegangen. Darüber hinaus sollten die beteiligten optischen Systeme die Bilder derselben Szene so ähnlich wie möglich numerisch darstellen. Dies betrifft einerseits die numerische Darstellung der Farben bzw. Pixel-Intensitäten und andererseits die optisch-geometrische Projektion der Szene auf den Sensor, der die Bildebene darstellt. Maßgeblich für eine identische numerische Darstellung von Farben ist sowohl

die spektrale Sensitivität der Sensoren als auch eine geeignete Steuerung der Belichtungszeit und der Verstärkung des CCD-Chip-Signals für alle beteiligten Kamerasysteme. Zudem kann durch eine pixelweise Transformation der Bilder die Gleichheit der Farbdarstellung optimiert werden. Die Projektion der Bilder auf dem Sensor wird durch die Lage der Bildebenen zueinander und die Abbildungsfehler der Objektive beeinflusst. Auf Basis der Kenntnis des Abbildungsverhaltens sowie der Position der Kameras zueinander ist es möglich, die Projektionsfehler zu reduzieren. Für eine optimale Lösung des Korrelationsproblems müssen die Abbildungsfehler der eingezogenen Bilder durch eine Rektifizierung minimiert werden. Das setzt eine Kalibrierung des Kamerasystems voraus.

Der Abschnitt 3.3 beschreibt ein Verfahren, das es ermöglicht, die Kameraelektronik so einzustellen, dass eine optimale numerische Darstellung der Farben bzw. Pixel-Intensitäten gewährleistet wird. Das Kapitel 3.4 geht auf die Kalibrierung des optischen Systems sowie der Rektifizierung der Bilddaten ein.

3.2.1 Belichtungsverfahren

Da sich das zu messende Objekt während des Messvorgangs relativ zum Träger des Messsystems bewegt, ist es notwendig, die Bilder aller beteiligten Kameras zeitgleich einzuziehen. Durch die Relativbewegung der Objekte hätte eine sequentielle Bilddatenerfassung einen zusätzlichen Versatz der korrespondierenden Bildpunkte zur Folge. Dieser zusätzliche Versatz würde die Suche nach den korrespondierenden Punktepaares erschweren. Die Forderung nach einem zeitgleichen Bildeinzug setzt voraus, dass der Bildeinzug durch die einzusetzende Kameragerätetechnik durch ein einziges Triggersignal gestartet werden kann.

Grundsätzlich sind zurzeit Digitalkameras mit zwei unterschiedlichen Belichtungsverfahren verfügbar. Einerseits Kameras, die das *Rolling-Shutter*-Verfahren nutzen. Andererseits Kameras, die auf dem *Global-Shutter*-Verfahren basieren. Beim Rolling-Shutter-Verfahren werden die Sensorpixel des CCD-Chips sequentiell während der Belichtung abgetastet. Die Abbildung 3.3 veranschaulicht das Prinzip der Sensorabtastung nach dem Rolling-Shutter-Verfahren.

Die Verwendung von Kameras, die das Rolling-Shutter-Verfahren nutzen, ist für die Aufnahme bewegter Szenen ungeeignet, da die Szenen abhängig von der Geschwindigkeit der Bewegung und der Abtastgeschwindigkeit verzerrt abgebildet werden. Die Abbildung 3.4 wie ein bewegtes Objekt durch den Rolling-Shutter-Bildeinzug verzerrt abgebildet wird. Ohne die Kenntnis der relativen Geschwindigkeit des betrachteten Objekts zur Kamera ist eine Entzerrung der im Rolling-Shutter-Verfahren aufgenommenen Bilder nur schwer möglich.

Im Global-Shutter-Verfahren werden alle sensitiven Elemente des Sensors-Chips zeitgleich ausgelesen. Die sensitiven Elemente der Kamera werden im Folgenden als Sensorpixel oder Pixel bezeichnet.

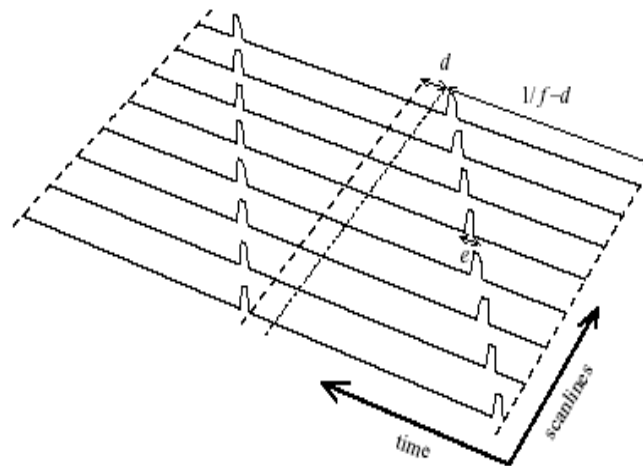


Bild 3.3: Sensorabtastung im Rolling-Shutter-Betrieb

Verzerrungen des Bildes, wie im Rolling-Shutter-Verfahren, treten hier nicht auf. Aus diesem Grund werden für die Erfassung der auszuwertenden Bilder Global-Shutter-Kameras verwendet.



Bild 3.4: Verzerrte Abbildung bewegter Objekte durch sequentielle Abtastung

3.3 Die Steuerung der Kameraelektronik

Das Ziel der Kamerasteuerung besteht darin, die Bilder unabhängig von den natürlich variierenden Lichtverhältnissen mit konstanter Qualität einzuziehen. Hierfür ist es zunächst notwendig, ein Gütekriterium zur Bewertung der Bildqualität zu definieren. Es wird eine Funktion gesucht, die aus den zweidimensionalen Datenfeldern der Bilddaten einen skalaren Qualitätswert berechnet. Zunächst werden Merkmale definiert, anhand derer die Qualität der aufgenommenen Bilder mit Hinsicht auf die weitere Bildauswertung beurteilt werden kann.

3.3.1 Die Bildqualität

Die Qualität eines Bildes kann als der Informationsgehalt des Bildes hinsichtlich der für den weiteren Prozess relevanten Merkmale bezeichnet werden. Im Rahmen dieser Bildauswertung werden abhängig vom Prozessschritt zum Teil widersprüchliche Forderungen an die Bildqualität gestellt.

Für die Extraktion der roten Ränder der Ladefläche ist der Kontrast der Ränder zum übrigen Bild von Interesse. Schon bei geringer Belichtung der aufgenommenen Szene erreichen die Bildbereiche, die den rot markierten Rand des Anhängers zeigen, den Bereich der Farbsättigung für diesen Farbton. Für eine Extraktion dieser Bildbereiche ist es vorteilhaft, wenn das Bild so belichtet wird, dass die Sättigung dieser Bildbereiche gerade erreicht wird. Eine intensivere Belichtung der Bilder würde keine Verbesserung des hier relevanten Informationsgehaltes bewirken. An dieser Stelle hätte eine weitere Verstärkung der Belichtung nur eine Verstärkung der Störungen in den übrigen Bildbereichen zur Folge.

Im Gegensatz dazu stellt die Textur die relevante Information für das Finden der korrespondierenden Punktpaare dar. Der Informationsgehalt einer Textur wird unter anderem durch den Kontrast, mit dem die Textur ausgeprägt ist, beeinflusst. An dieser Stelle ist eine Belichtung, die einen hohen Kontrast in dem Bildbereich, der das Ladegut zeigt, gefordert. Details über den Zusammenhang zwischen der Textur, dem Kontrast und dem sich daraus ergebenden Informationsgehalt für das Korrelationsproblem werden in Kapitel 5.2 erläutert. Die Belichtung wird hier als optimal angesehen, wenn sie mit hinreichender Sicherheit den notwendigen Kontrast im Bereich des Ladegutes bietet.

3.3.2 Ein Gütefunktional zur Bewertung der Bildqualität

Für eine objektiv quantitative Bewertung der Bildqualität wird das in diesem Abschnitt beschriebene Gütefunktional verwendet. Mit Hinsicht auf die Berechnungszeit ist es unter anderem das Ziel, das Gütefunktional möglichst einfach zu formulieren. Hierfür erweist sich die gewichtete Summe aus dem mittleren Helligkeitswert B_{mean} , dem maximalen Helligkeitswert B_{max} und einem den Clippingeffekt beschreibenden Wert B_{clip} als geeignet. Die Bildqualität Q berechnet sich nach Gleichung 3.1.

$$Q = aB_{mean} + bB_{max} + cB_{clip} \quad (3.1)$$

Hierbei gewichten die Koeffizienten a, b, c den Einfluss der jeweils zugehörigen Bildmerkmale. Der mittlere Helligkeitswert B_{mean} entspricht dem über alle Pixel und Farbkanäle gebildeten Mittelwert der auftretenden Intensitäten eines Bildes. Er ist daher ein Maß für die Bildhelligkeit. Wird eine gleichmäßige Verteilung der Intensitätswerte angenommen, gibt B_{mean} einen guten Richtwert für die Bildqualität. Der maximale Helligkeitswert B_{max} eines Bildes entspricht dem Mittelwert der Intensitätsmaxima der beteiligten Farbkanäle und berechnet sich gemäß Gleichung 3.2.

$$B_{max} = \frac{\max(I_R) + \max(I_G) + \max(I_B)}{3} \quad (3.2)$$

I_R bezeichnet die Menge der auftretenden Intensitäten im Rot-Layer, I_G die Menge der Intensitäten im Grün-Layer und I_B die Menge der auftretenden Intensitäten im Blau-Layer des betrachteten Bildes. Der Wert B_{max} gibt Aufschluss über die Ausnutzung des Wertebereiches mit Hinsicht auf die Obergrenze des Wertebereiches. Der Wert B_{clip} berechnet sich nach Gleichung 3.3

$$B_{clip} = \frac{H_{255}}{H_{255,max}} \quad (3.3)$$

$H(255)$ ist die Anzahl der Pixel, deren Intensitätswert die obere Grenze des Wertebereiches erreicht haben. $H(255)_{max}$ gibt die maximal tolerierbare Menge an Pixeln an, deren Intensitätswert die obere Grenze des Wertebereiches erreicht haben. Mit dem Quotienten aus der Anzahl tatsächlich gesättigter Pixel und der Anzahl maximal erlaubter gesättigter Pixel erhält man ein Maß für den Grad der Überbelichtung des Bildes.

3.3.3 Die Regelung der Bildqualität

Um bei natürlich variierenden Lichtverhältnissen eine gleichbleibende Qualität der eingezeichneten Bilder zu gewährleisten, wird über einen Regelalgorithmus sowohl die Belichtungszeit als auch die Verstärkung der CCD-Sensoren eingestellt. Die Abbildung 3.5 zeigt ein Schaubild der verwendeten Regelungsstruktur. Das auf den CCD-Sensor fallende Licht generiert abhängig von der Belichtungszeit $t_{exp}(exposure)$ und dem Blendenwert k ein analoges Videosignal. Dieses wird mit dem gewählten Wert V_{ccd} verstärkt und anschließend durch den A/D-Wandler digitalisiert. Das Gütefunktional berechnet aus dem Bild einen Qualitätswert Q , der die Regelgröße darstellt. Das Übertragungsverhalten der Sensorchips von der Beleuchtungsintensität zum wiedergegebenen Bildhelligkeit, so wie das Abbildungsverhalten des Bild-Gütefunktional sind deutlich nichtlinear. Um diese Nichtlinearität auszugleichen, wird in einem Stellalgorithmus die inverse Kennlinie des Abbildungsverhaltens der Kameras und des Bild-Gütefunktional auf die Reglerausgangsgröße y_R angewendet. Zusätzlich berechnet der Stellalgorithmus aus Reglerausgangsgröße y_R die neuen Werte für die Belichtungszeit t_{exp} und die Verstärkung V_{ccd} des CCD-Signals. Die folgenden Aufnahmen werden dann mit diesen Kameraeinstellungen eingezeichnet. Da so die Bilder beider beteiligten Kameras mit dem gleichen Gütefunktional hinsichtlich ihrer Qualität bemessen werden und darüberhinaus beide Kameras getrennt auf den selben Bildqualitäts-Sollwert geregelt werden, kann im eingeschwungenen Zustand von annähernd gleicher Qualität der bereitgestellten Bilder ausgegangen werden.

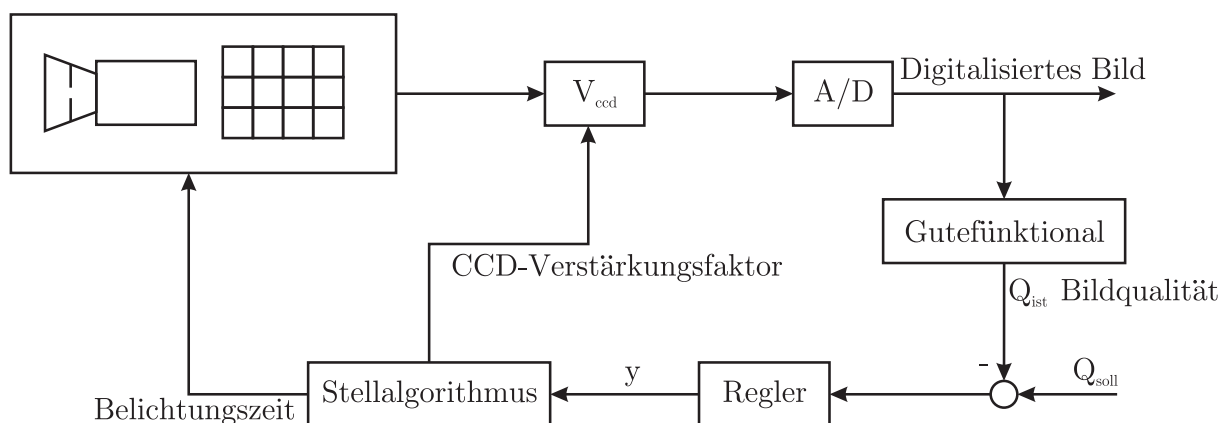


Bild 3.5: Regelungsstruktur

3.3.4 Die Qualität der eingezogenen Bildpaare

Die Kameras sind so angeordnet, dass beide Kameras annähernd die gleiche Szene abbilden. Im Idealfall würden die eingezogenen Bilder identische Intensitätsverteilungen der einzelnen Farbkomponenten aufweisen.

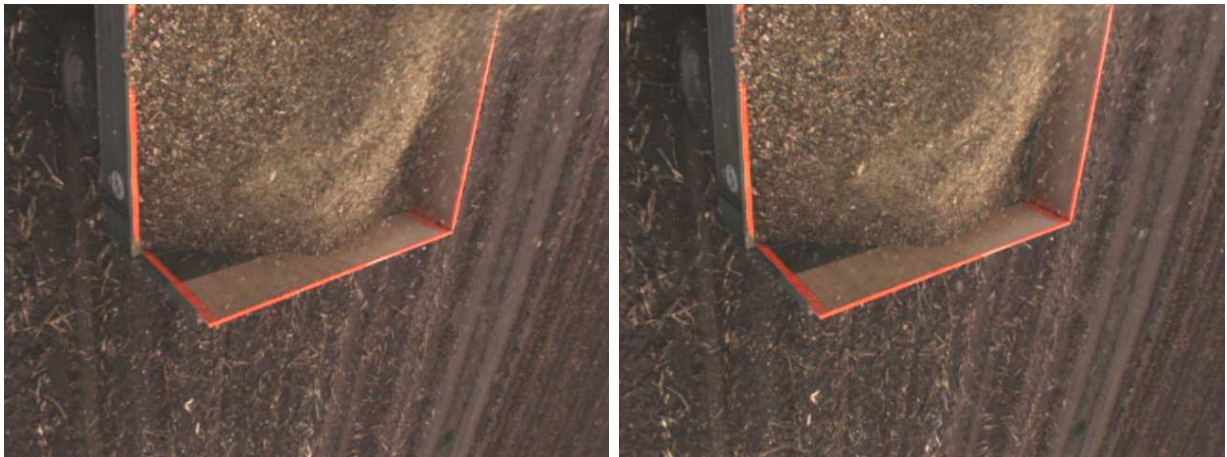


Bild 3.6: Eingezogene Bilder

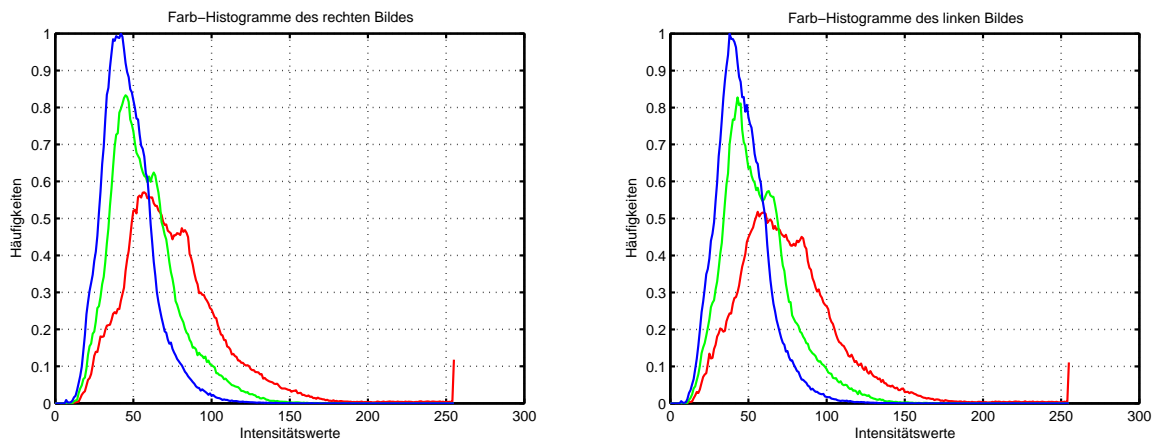


Bild 3.7: Farbkanal-Histogramme der eingezogenen Bilder

Die Abbildungen 3.6 zeigen einen Bilddatensatz, der zur Auswertung des Füllstandes herangezogen wird. Die Abbildungen 3.7 zeigen die zugehörigen Farbkanal-Histogramme. Hierbei ist zu beachten, dass die Häufigkeiten auf die Anzahl der Pixel in einem Bild normiert wurden. Es wird deutlich, dass einerseits sowohl die Verläufe der Histogramme einer Farbe als auch die Verhältnisse der Histogramme der einzelnen Farben zueinander annähernd identisch

sind. Mit Hilfe dieser Regelung ist es also möglich, die Kameras unabhängig voneinander so einzustellen, dass die Gleichheit der numerischen Repräsentation der abgebildeten Farben gewährleistet ist.

3.4 Die Kalibrierung des Epipolarsystems

Für die Füllstandsmessung wird ein Stereokamerasystem verwendet. Die Kerninformation für die stereoskopische Bildauswertung ist die Disparität. Ein wesentliches Problem stellt somit die Suche nach dem korrespondierenden Punkt P_R im rechten Bild zu einem Referenzpunkt P_L im linken Bild dar, was im Allgemeinen als Korrelationsproblem oder Pointmatching Problem bezeichnet wird. Für die Lösung des Korrelationsproblems sowie die Rekonstruktion von Objektpunkten im Rahmen der Stereovisionsanalyse ist die Kenntnis des Abbildungsverhalten des verwendeten Epipolarsystems von grundlegender Bedeutung. Dies betrifft sowohl die Anordnung der Kameras zueinander als auch das Abbildungsverhalten der einzelnen optischen Systeme. Die Lage der Kameras in den sechs Freiheitsgraden im Weltkoordinatensystem werden als die extrinsischen Parameter des Epipolarsystems bezeichnet. Die intrinsischen Parameter beschreiben das Abbildungsverhalten einer Kamera. Durch eine Kalibrierung des Epipolarsystems werden diese Parameter festgestellt [16, 7, 20, 21, 22]. Die Kalibrierung des Kamerasystems wird anhand eines bekannten Testmusters durchgeführt. Hier kommt das Matlab-Werkzeug *A Camera Calibration Toolbox for Matlab* [6] zum Einsatz. Ein bekanntes Referenzmuster wird in unterschiedlichen Entfernungen zur Kamera und in verschiedenen Perspektiven aufgenommen und ausgewertet. So wird im Rahmen einer Offline-Kalibrierung der Satz der extrinsischen und intrinsischen Parameter bestimmt.

Die Verwendung eines so ermittelten Parametersatzes setzt voraus, dass sich das Kamerasystem während der Messung nicht verändert. Dies betrifft einerseits die Anordnung der Kameras zueinander, was durch eine hinreichend steife Montage der Kameras erreicht wird. Andererseits darf die Einstellung der optischen Systeme nicht verändert werden. Aus diesem Grund wird von der Verwendung automatisch fokussierender Objektive abgesehen. Es kommen Standard-Weitwinkel-Objektive mit manuell einstellbarer und fixierbarer Fokussierung sowie Blendenöffnung zum Einsatz.

Der folgende Abschnitt 3.4.1 geht kurz auf die grundsätzliche Vorgehensweise der Kamerakalibrierung ein. Eine detaillierte Betrachtung der Kalibrierungsverfahren ist nicht Gegenstand dieser Arbeit. An dieser Stelle wird auf die einschlägige Literatur verwiesen.

3.4.1 Vorstellung der unterschiedlichen Kalibrierverfahren

Die Methoden zur Kalibrierung können in zwei Kategorien unterteilt werden:

- **Photogrammetrische Kalibrierung:** Die Kalibrierung wird durchgeführt, indem man einen Kalibrierungsgegenstand beobachtet, dessen Geometrie im dreidimensionalen Raum mit sehr guter Präzision bekannt ist. Diese Art der Kalibrierung kann sehr leistungsfähig erfolgen [22]. Der Kalibrierungsgegenstand besteht normalerweise aus zwei oder drei Flächen, die orthogonal zueinander sind. Diese Methoden erfordern einen kostspieligen Kalibrierungsapparat und eine durchdachte Einstellung.
- **Selbst-Kalibrierung:** Techniken dieser Kategorie benutzen keinen Kalibrierungsgegenstand. Beim Verschieben einer Kamera in einer statischen Szene stellt die Starrheit der Szene zwingend zwei der internen Kameraparametern zur Verfügung [26]. Wenn folglich Bilder durch die gleiche Kamera mit örtlich festgelegten internen Parametern aufgenommen werden, genügen Korrespondenzen zwischen drei Bildern, um die internen und externen Parameter zu gewinnen. Da die Anzahl der zu schätzenden Parameter hoch ist, können nicht immer zuverlässige Resultate gewonnen werden.

Es sei angemerkt, dass noch weitere Verfahren [18, 27] existieren, die hier aber nicht weiter betrachtet wurden. Das verwendete Kalibrierungsverfahren nach Zhang erzielt mit einfachen Mitteln die Modellparameter des Kamerasystems in hinreichender Genauigkeit [20].

3.4.2 Kalibrierung nach Zhang

Dieses Verfahren benutzt ein planares Schachbrettmuster (vgl. Abbildung 3.8) als Kalibrierobjekt, wobei die Berührungspunkte der einzelnen Schachfelder als Kalibriermarken genutzt werden [20, 21]. Zur Bestimmung der intrinsischen und extrinsischen Parameter sind mindestens zwei Bilder des Musters mit unterschiedlichen Orientierungen nötig, wobei diese Orientierungen nicht bekannt sein müssen.

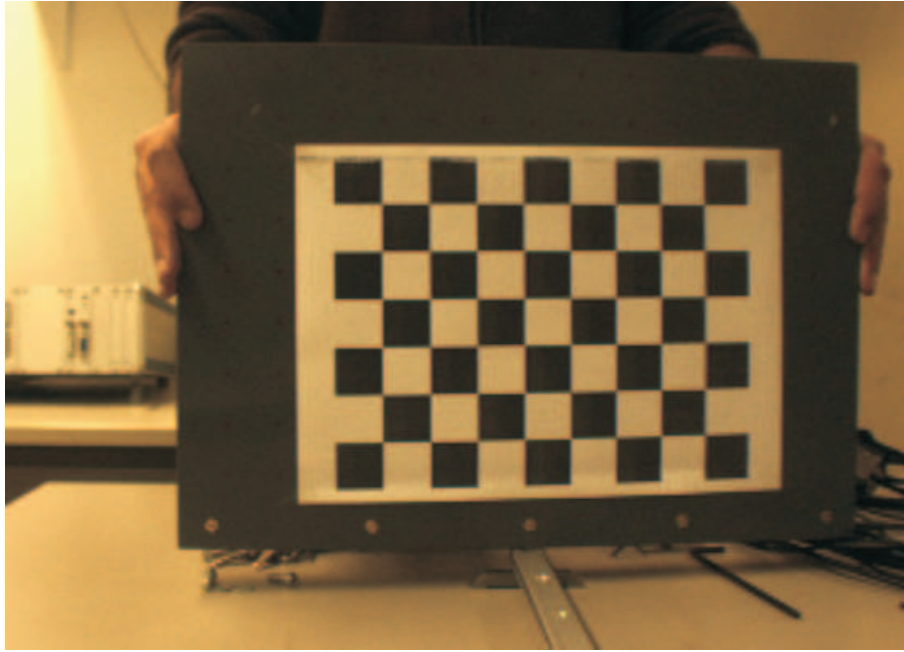


Bild 3.8: Kalibrierungsmuster

Die Bestimmung der Kameraparameter mit diesem Verfahren lässt sich in drei Schritte unterteilen:

1. Schätzung der Homographiematrix \mathbf{H}
2. Bestimmung der Kameraparameter aus der Homographiematrix \mathbf{H}
3. Modellierung der Verzerrung und Optimierung der Gesamtlösung

3.4.3 Abschätzung der Homographiematrix \mathbf{H}

Um die Homographiematrix zu bestimmen, wird in einer beliebigen Entfernung zum Kamerasystem ein planares Kalibriermuster aufgenommen. Diese Entfernung wird als $Z = 0$ bezeichnet. Bezeichnet man die i -te Spalte der Rotationsmatrix \mathbf{R} mit R_i , so ergibt sich die Abbildung P' des Objektpunktes P nach Gleichung 3.4.

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{A}[\mathbf{R} \vec{t}] \begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = \mathbf{A}[R_1 R_2 R_3 \vec{t}] \begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = \mathbf{A}[R_1 R_2 \vec{t}] \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \quad (3.4)$$

Da die Entfernung des Kalibrierusters zum Kamerasystem noch nicht bestimmt ist, muss der entfernungsabhängigen zentrisch gestreckten Abbildung durch den skalaren Faktor s Rechnung getragen werden. Die Homographie H für P und P' ergibt sich entsprechend Gleichung 2.32 nach

$$sP' = \mathbf{H}_{\mathbf{P}\mathbf{P}'} P \quad (3.5)$$

$$\mathbf{H}_{\mathbf{P}\mathbf{P}'} = \mathbf{A}[R_1 R_2 \vec{t}'] \quad (3.6)$$

Entsprechend der Anzahl der unbekannten Größen werden nun hinreichend viele Bildpunkte P_{ij} aus unterschiedlichen Aufnahmen des Kalibrierusters ausgewählt, um das Gleichungssystem 3.5 zu lösen. Im Allgemeinen wird ein überbestimmtes Gleichungssystem erzeugt, um durch eine Minimierung des quadratischen Fehlers nach Gleichung 3.7 eine optimale Lösung zu erhalten.

$$\sum_{j=1}^n \sum_{i=1}^m \|P'_{ij} - \hat{P}'(H, P_{ij})\|^2 \quad (3.7)$$

Die Minimierung dieses nichtlinearen Problems kann z.B. mit dem Levenberg-Marquardt-Algorithmus C erfolgen.

3.4.4 Bestimmung der intrinsischen und extrinsischen Parameter

Gemäß Gleichung 2.33 beinhaltet die Homographiematrix $\mathbf{H}_{\mathbf{L}\mathbf{R}}$ für ein Stereokamerasystem die Abbildungsmatrizen $\mathbf{A}_{\mathbf{L}}$ und $\mathbf{A}_{\mathbf{R}}$ der beiden Kameras sowie die Rotationsmatrix $\mathbf{R}_{\mathbf{R}}$ und den Translationsvektor \vec{t}_R , welche die Lage der Kameras zueinander beschreiben. Wie in Kapitel 2.1 beschrieben, sind somit sowohl die extrinsischen als auch die intrinsischen Kameraparameter in der Homographiematrix enthalten. Mit der bereits bestimmten Homographiematrix kann gemäß Gleichung 2.33 demnach wieder ein nichtlineares Gleichungssystem aufgestellt werden. Die so gefundenen Ergebnisse können wiederum durch Minimierung des quadratischen Fehlers (vgl. Gleichung 3.8) zwischen dem abgebildeten Punkt P' und dem rekonstruierten Punkt $\hat{P}'(\mathbf{A}, \mathbf{R}, \vec{t}, P)$ gefunden werden.

$$\sum_{j=1}^n \sum_{i=1}^m \|P'_{ij} - \hat{P}'(A, R_j, \vec{t}_j, P_i)\|^2 \quad (3.8)$$

3.4.5 Modellierung der Verzeichnung und Optimierung der Gesamtlösung

Um die Parameter für die Linsenverzeichnung zu finden, werden wiederum die Ergebnisse des vorgehenden Schrittes aus Kapitel 3.4.4 als Anfangslösung verwendet. Auch hier wird der quadratische Fehler des mit dem Modell der Linsenverzeichnung rekonstruierten Fehlers minimiert. Durch Minimierung der Gleichung 3.9 werden die Parameter k_1 und k_2 der radialen Verzeichnung bestimmt.

$$\sum_{j=1}^n \sum_{i=1}^m \|P'_{ij} - \hat{P}'(A, k_1, k_2, R_j, \vec{t}_j, P_i)\|^2 \quad (3.9)$$

Dabei ist $\hat{P}'(A, k_1, k_2, R_j, \vec{t}_j, P_i)$ als Projektion des Punktes P_i in Bild j nach 3.5 zu verstehen. Als Startwerte für A und $R_j, \vec{t}_j | j = 1, \dots, n$ dienen die geschätzten und optimierten Werte aus dem vorherigen Schritt. Die Parameter k_1 und k_2 werden zu Beginn der Optimierung gleich 0 gesetzt.

3.4.6 Kameraparameter des Messsystems

Die Kalibrierung des verwendeten Kamerasystems hat den in diesem Abschnitt beschriebenen Parametersatz ergeben. Diese Werte werden für alle weiteren Betrachtungen herangezogen.

Parameter	Wert	3σ
Brennweite c	629.667	1.647
Skalierungsfaktor m	0.998	0.039
Hauptpunkt in u_0	336.278	3.817
Hauptpunkt in v_0	235.219	2.414
Radiale Verzeichnung k_1	-0.22012	0.007
Radiale Verzeichnung k_2	0.20887	0.019
Tangentiale Verzeichnung p_1	-0.00058	0.00058
Tangentiale Verzeichnung p_2	0.00028	0.00091

Tabelle 3.1: Intrinsische Parameter: Linke Kamera

Parameter	Wert	3σ
Brennweite c	629.67355	1.64709
Skalierungsfaktor m	0.99794	0.040404
Hauptpunkt in u_0	315.56911	3.60792
Hauptpunkt in v_0	228.47697	2.19536
Radiale Verzerrung k_1	-0.21990	0.00870
Radiale Verzerrung k_2	0.22697	0.04262
Tangentiale Verzerrung p_1	0.00092	0.00059
Tangentiale Verzerrung p_2	0.00076	0.00086

Tabelle 3.2: Intrinsische Parameter: Rechte Kamera

Parameter	Wert	3σ
Translation in x t_x	147.83770	0.11577
Translation in y t_y	-0.72179	0.08155
Translation in z t_z	2.17891	0.93059
Rotation um x θ_x	-0.00378	0.00228
Rotation um y θ_y	-0.00325	0.00488
Rotation um z θ_z	-0.01606	0.00038

Tabelle 3.3: Extrinsische Parameter

An den intrinsischen Parametern in den Tabellen 3.1 und 3.2 der rechten und linken Kamera, die die tangentiale Verzeichnung beschreiben, wird deutlich, dass der Einfluss der tangentialen Verzeichnung deutlich geringer ist als der Einfluss der radialen Verzeichnung. Dies bestätigt die vorhergehende Annahme, dass der Einfluss der tangentialen Verzeichnung vernachlässigbar klein ist. Darüber ist die Standardabweichung dieser Werte so groß, dass die Verwendbarkeit fraglich ist. Die Verzeichnungsvektoren lassen sich mit der Funktion `cvInitUndistortMap()` der *Intel OpenCV Bibliothek* [8] für jedes Pixel berechnen. Die Pixelverschiebungen werden in den Abbildungen 3.9 bis 3.14 durch Vektoren verdeutlicht. Die ellipsenförmigen Äquipotentiallinien bezeichnen die Bereiche, in denen ein konstanter Versatz der Pixel auftritt. Das Kreuz \times ist der durch den Adressbereich definierte Mittelpunkt des Bildes. \circ gibt den Hauptpunkt wieder.

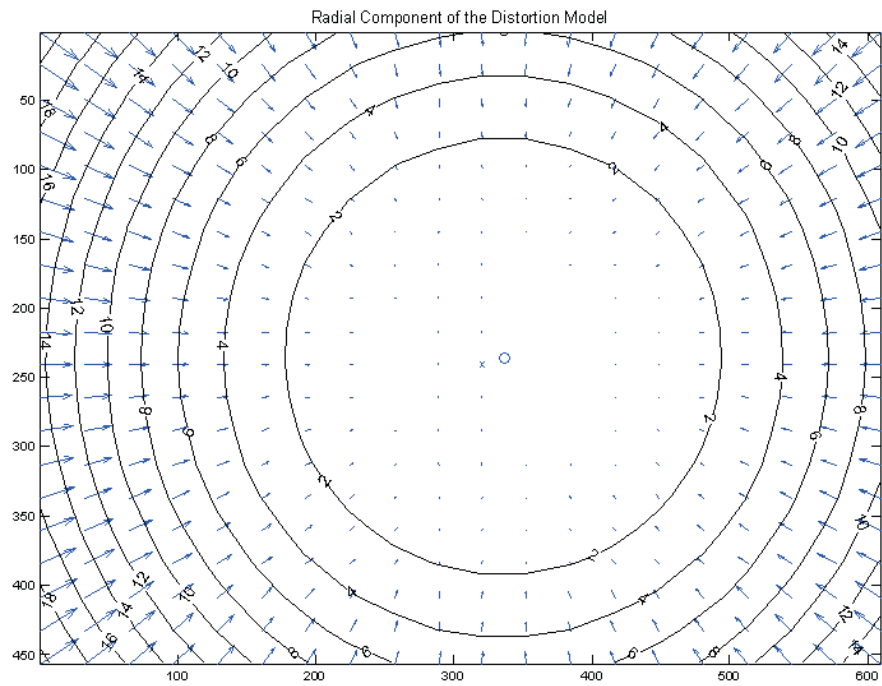


Bild 3.9: Radiale Verzerrung der linken Kamera

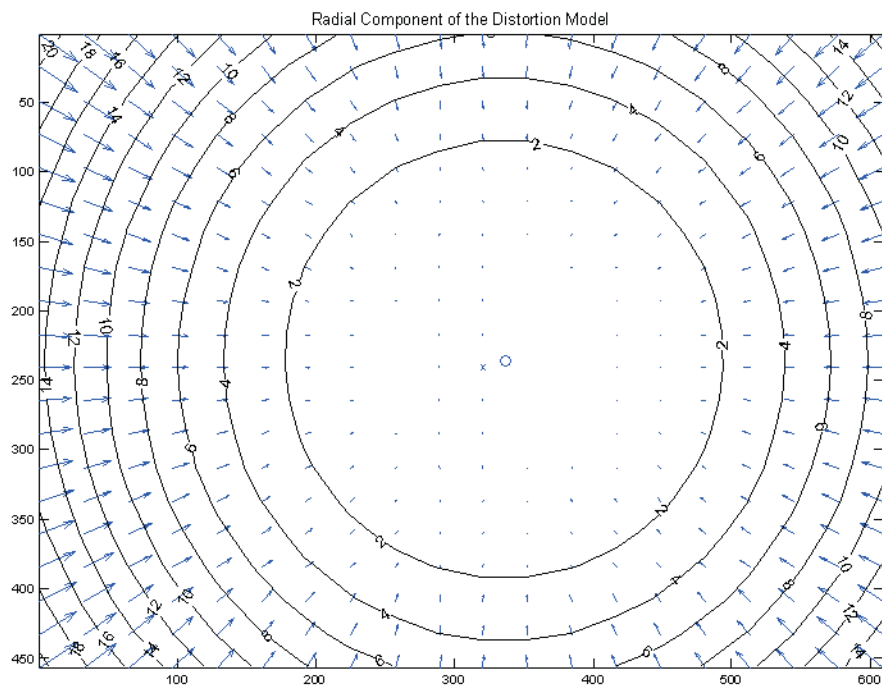


Bild 3.10: Radiale Verzerrung der rechten Kamera

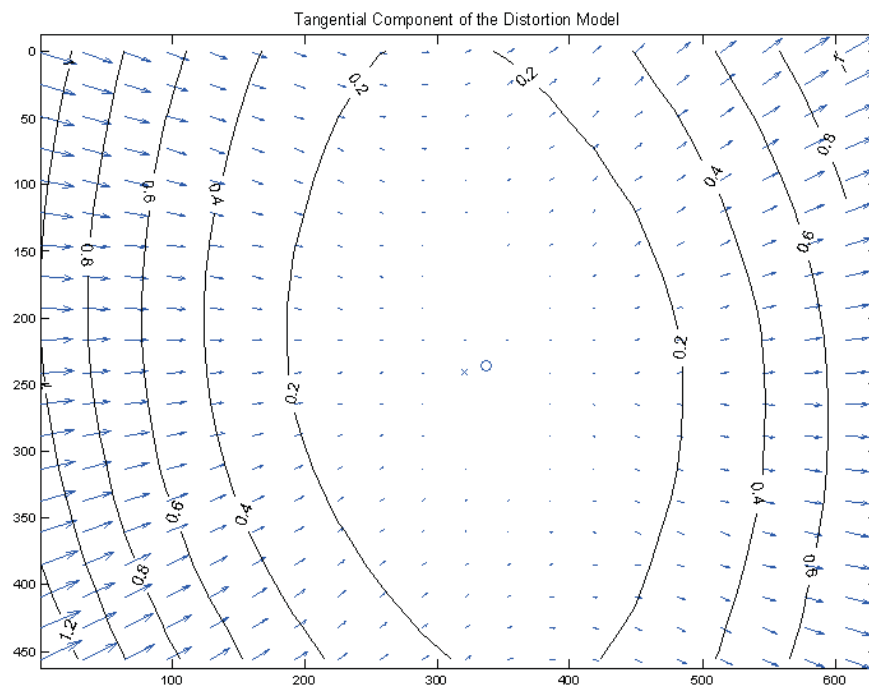


Bild 3.11: Tangentiale Verzerrung der linken Kamera

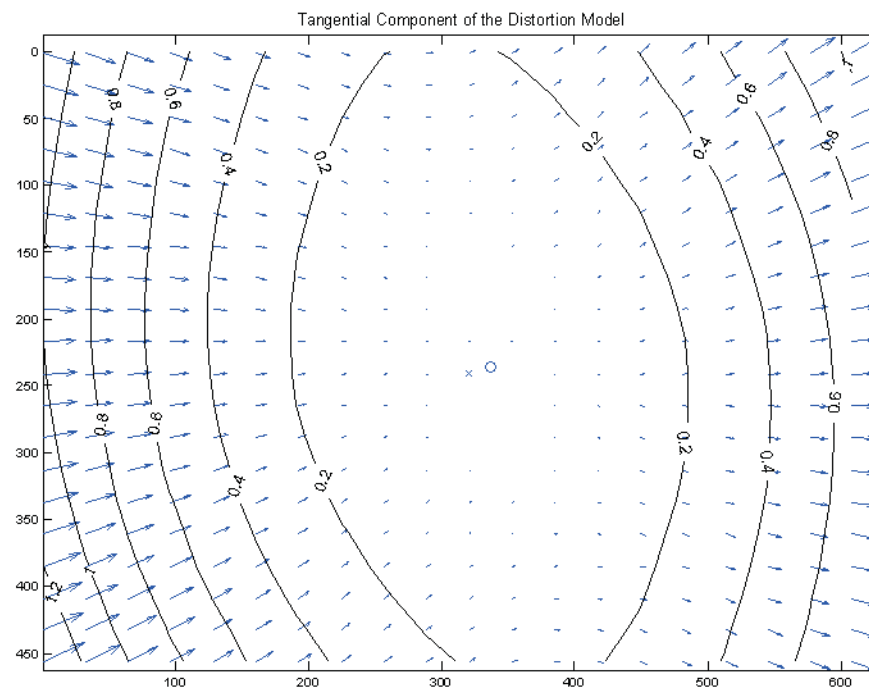


Bild 3.12: Tangentiale Verzerrung der rechten Kamera

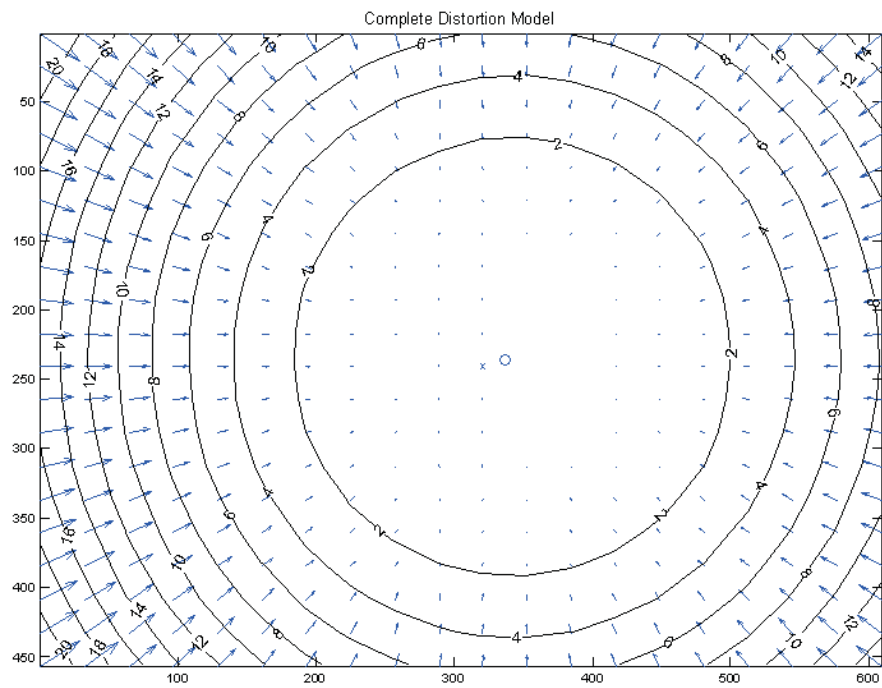


Bild 3.13: Gesamtverzerrung der linken Kamera

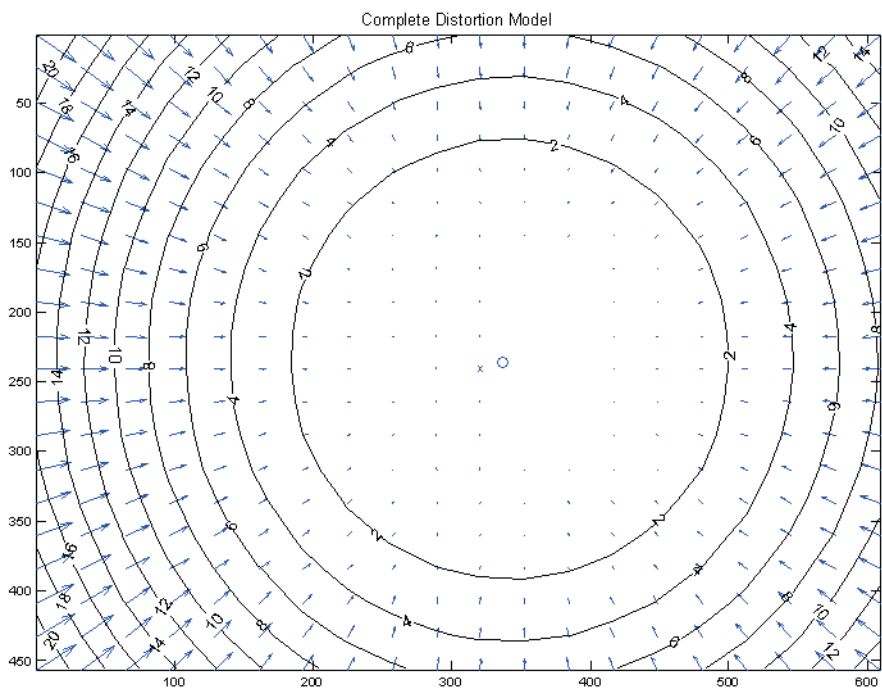


Bild 3.14: Gesamtverzerrung der rechten Kamera

3.5 Die Optimierung der eingezogenen Bilder

Neben einer optimalen Einstellung der Kamera-Hardware werden die eingezogenen Bilddaten im Rahmen der Bilddatenakquisition durch eine geeignete Software modifiziert. Die wichtigste Modifikation der Bilddaten kompensiert den Versatz der Kameras zueinander sowie die Linsenverzeichnung der Kameras. Diese Modifikation wird im Allgemeinen als *Rektifizierung* der Bilder bezeichnet. Die Rektifizierung transformiert die Bilder so, dass sie der Abbildung durch ein achsparalleles Stereokamerasystem mit Kameras, die einen identischen Satz intrinsischer Parameter aufweisen, hinreichend nahe kommen.

3.5.1 Rektifizierung

Es gibt zwei Methoden der Rektifizierung. Die erste Variante beruht auf der Kenntnis der intrinsischen und extrinsischen Kameraparameter. Sie wird auch *kalibrierte Rektifizierung* oder *geometrische Rektifizierung* [53] genannt. Die zweite Variante wird unter Zuhilfenahme der Fundamentalmatrix des Epipolarsystems durchgeführt. Dieses Verfahren bietet den Vorteil, dass der vollständige Satz der Kameraparameter nicht explizit bekannt sein muss. Da im vorliegenden Fall aufgrund der Kalibrierung des Kamerasystems der vollständige Satz der Modellparameter bekannt ist, wird in dieser Arbeit die geometrische Rektifizierung angewendet. Im folgenden Abschnitt wird die Rektifizierung mit kalibrierten Kameras näher erläutert.

Das Prinzip der Rektifizierung sieht einerseits eine Kompensation der Linsenverzeichnung andererseits eine Lagekorrektur der Bildebenen der Kameras vor. Die Kompensation der Linsenverzeichnung geschieht dabei durch die Anwendung der inversen Verzeichnungsfunktion auf jedes Pixel (vgl. Gleichungen 2.10, 2.12, 2.13). Die Lage der Bildebenen wird zunächst durch eine Rotation der Bilder um ihren Hauptpunkt mit dem negativen Rotationsvektor $\vec{\theta}$ korrigiert. Abschließend werden die Bilder so verschoben, dass nur ein translativer Versatz entlang der X -Achse verbleibt. Hierbei bleibt zu berücksichtigen, dass eine Translation entlang der Z -Achse einer Kamera eine Streckung oder Stauchung der Abbildung zur Folge hat, die dementsprechend in der Abbildungsmatrix 2.8 berücksichtigt werden muss. Es ist naheliegend, dass für die rektifizierten Bilder ein neuer Satz intrinsischer und extrinsischer Parameter gilt. Der neue Satz der Kameraparameter definiert sich weitestgehend aus den Anforderungen, die an die rektifizierten Bilder gestellt werden. Demnach müssen die Parameter eines ideal rektifizierten Epipolarsystems wie in Tabelle 3.4 lauten.

Parameter	Wert
Translation $\overrightarrow{t_{rect}}$	$\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}$
Rotation $\overrightarrow{\theta_{rect}}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$
Koeffizienten der radialen Verzeichnung k_i	$[0 \ 0]$
Koeffizienten der tangentialen Verzeichnung p_i	$[0 \ 0]$

Tabelle 3.4: Rektifizierte Kameraparameter

Hierbei sei angemerkt, dass die Koeffizienten für die radiale und tangentiale Verzeichnung einheitlich für alle beteiligten Kameras gelten. Der verbleibende zu bestimmende Parameter ist der translativ Versatz der rektifizierten Bildebenen entlang der X -Achse. Somit ergeben sich neue Projektionsmatrizen $\tilde{\mathbf{P}}_l$ und $\tilde{\mathbf{P}}_r$ für die rektifizierte Abbildung nach Gleichung 3.10 und 3.11

$$\tilde{\mathbf{P}}_l = \mathbf{A}[(\mathbf{R}\mathbf{R}_L^{-1})c_L] \quad (3.10)$$

$$\tilde{\mathbf{P}}_r = \mathbf{A}[(\mathbf{R}\mathbf{R}_R^{-1})c_R] \quad (3.11)$$

Die negativen Rotationsmatrizen \mathbf{R}_L und \mathbf{R}_R beschreiben die Transformation der Bildebenen in das Weltkoordinatensystem. Die Rotationsmatrix \mathbf{R} , die den Bildern beider Kameras entspricht, beschreibt die Rotation in das Zielkoordinatensystem. Ebenso transformiert die Anwendung der negativen Kamerakonstanten c_L und c_R der linken und rechten Kamera die Bilder in eine Abbildung durch Kameras mit der einheitlichen Kamerakonstante 1.

Bilineare Interpolation

Im Allgemeinen wird durch die Rotation sowie den Ausgleich der Linsenverzeichnung die Koordinate eines Pixels als Gleitkommawert abgebildet. Um nun die transformierten Pixel dennoch auf dem diskreten Raster der Bildebene darstellen zu können, ohne dass das Bild erneut verzerrt erscheint, wird der Intensitäts- oder Farbwert I_{uv} eines Pixels an der Stelle u , v durch eine Bilineare-Interpolation auf Basis der Farbwerte der vier nächsten Pixel ermittelt. Die Abbildung 3.15 veranschaulicht die Problematik der Rotation von Bildern.

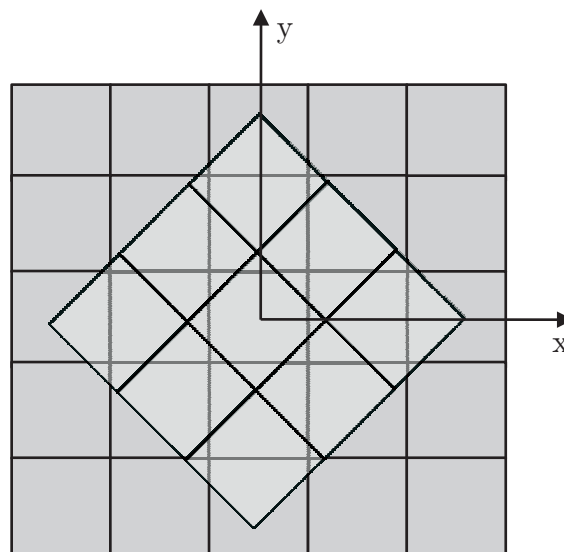


Bild 3.15: Rotation eines Bildes

Es wird deutlich, dass die hier um die Z -Achse gedrehten Pixel nicht genau auf das Raster der Bildebene abgebildet werden. Um nun einen geeigneten Farbwert I_{uv} des Pixels P_{uv} an der Stelle u,v nach der Transformation ermitteln zu können, werden die Farbgradienten zwischen den Pixeln des transformierten Bildes verwendet.

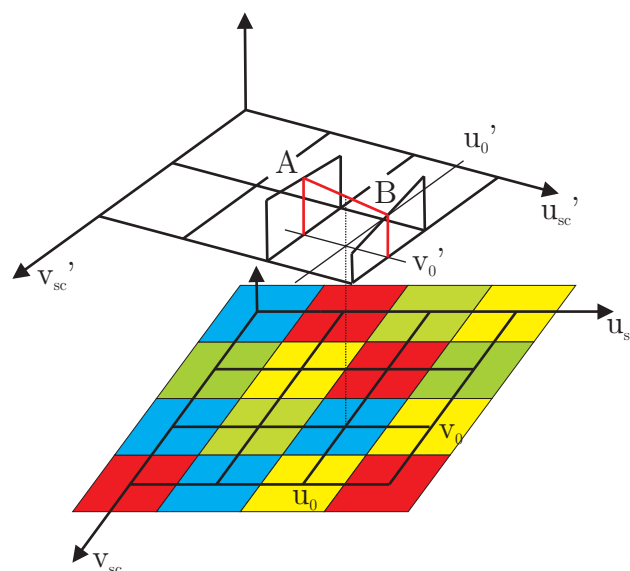


Bild 3.16: Geometrische Deutung der bilinearen Interpolation

Die Abbildung 3.16 zeigt die geometrische Interpretation der bilinearen Interpolation. An dieser Stelle wird der Farbwert eines um die W -Achse gedrehten Pixels gesucht. Hier soll das

obere, gedrehte Raster auf das untere Raster der Bildebenen übertragen werden. Hierbei sind die Farbwerte bzw. die Intensitäten der einzelnen Pixel auf der W -Achse aufgetragen. Die Schnittpunkte des Gitters repräsentieren die Zentren der Pixel. Soll nun das transformierte Bild an die Stelle u_0, v_0 der Bildebenen übertragen werden, so wählt man die vier zu diesem Punkt nächsten Pixel des transformierten Bildes. Die Positionen dieser vier Pixel können mit $[u'_0, v'_0]$, $[u'_1, v'_0]$, $[u'_0, v'_1]$ und $[u'_1, v'_1]$ bezeichnet werden. Entsprechend lauten die Intensitätswerte des transformierten Bildes an diesen Stellen I_{00} , I_{10} , I_{01} und I_{11} . Durch die Interpolation entlang einer gewählten Achse erhält man zwei neue Intensitätswerte A und B nach Gleichung 3.12 und 3.13

$$A = \frac{I_{01} - I_{00}}{v'_1 - v'_0} \tilde{v}_0 + I_{00} \quad (3.12)$$

$$B = \frac{I_{11} - I_{10}}{v'_1 - v'_0} \tilde{v}_0 + I_{10} \quad (3.13)$$

\tilde{v}_0 bezeichnet die U -Koordinate der Durchdringung des Lots (blau eingezeichnet) auf die Bildebenen an der Stelle $[u_0, v_0]$ mit der transformierten Bildebene. In Abbildung 3.16 sind die interpolierten Intensitäten A und B als rote Linien eingezeichnet. Der Intensitätswert $I_{u_0 v_0}$, der in die Bildebene übertragen wird, ergibt sich nun durch eine Interpolation zwischen den zuvor gefundenen Intensitäten A und B . Die Intensität $I_{u_0 v_0}$ an der Stelle u_0, v_0 ergibt sich dann nach Gleichung 3.14.

$$I_{u_0 v_0} = \frac{B - A}{u_1 - u_0} \tilde{u}_0 + A \quad (3.14)$$

Hierbei bezeichnet \tilde{u}_0 analog zu Gleichung 3.12 und 3.13 die U -Koordinate der Durchdringung des Lots auf die Bildebenen an der Stelle $[u_0, v_0]$ durch die Ebenen der transformierten Bildebene.

Implementierung der Rektifizierung

Wie schon in Kapitel 2 erwähnt, wird das Weltkoordinatensystem S_0 in das Koordinatensystem S_{KL} der linken Kamera gelegt. Auch hier zeigt sich, dass diese Wahl sinnvoll ist und die Transformation der Bilder deutlich vereinfacht. So muss nun in den Bildern der linken Kamera nur die Linsenverzeichnung ausgeglichen werden, während die Bilder der rechten Kamera neben dem Ausgleich der Linsenverzeichnung in die Bildebene der linken Kamera transformiert werden müssen. Die Einträge der Abbildungsmatrizen für die rektifizierten

Bilder der linken und rechten Kamera ergeben sich nach den Gleichungen 3.15 und 3.16.

$$\tilde{\mathbf{P}}_{\mathbf{L}} = -\delta\vec{P}_L(P_L) \quad (3.15)$$

$$\tilde{\mathbf{P}}_{\mathbf{R}} = \mathbf{A}[\mathbf{R}_{\mathbf{R}}^{-1} - c_r] - \delta\vec{P}_R(P_R) \quad (3.16)$$

$$\text{mit} \quad \delta\vec{P}_i = \begin{pmatrix} \delta u(u) \\ \delta v(v) \end{pmatrix} \quad (3.17)$$

In Gleichung 3.17 bezeichnen $\delta u(u)$ und $\delta v(v)$ den Versatz der Linsenverzeichnung. Entsprechend der Gleichungen 2.10, 2.12 und 2.13 Komponenten der Korrekturvektors Funktionen des Pixelortes u, v . Unter der Voraussetzung, dass das Kamerasystem während der Messung nicht verändert wird, kann davon ausgegangen werden, dass alle eingezogenen Bilder einer Kamera mit identischen Parametern transformiert werden können. Mit Hinsicht auf eine laufzeitoptimierte Implementierung werden zunächst auf der Basis der in Kapitel 3.4.4 aufgeführten Werte für jede Kamera Korrekturmatriizen M_K erstellt. Diese Korrekturmatriizen enthalten für jedes Pixel den Vektor um den das Pixel in der Bildebene verschoben werden muss, um die Abbildungsfehler des Kamerasystems auszugleichen. Die Abbildung 3.17 zeigt die Rektifizierungsmatrix für die rechte Kamera des Kamerasystems.

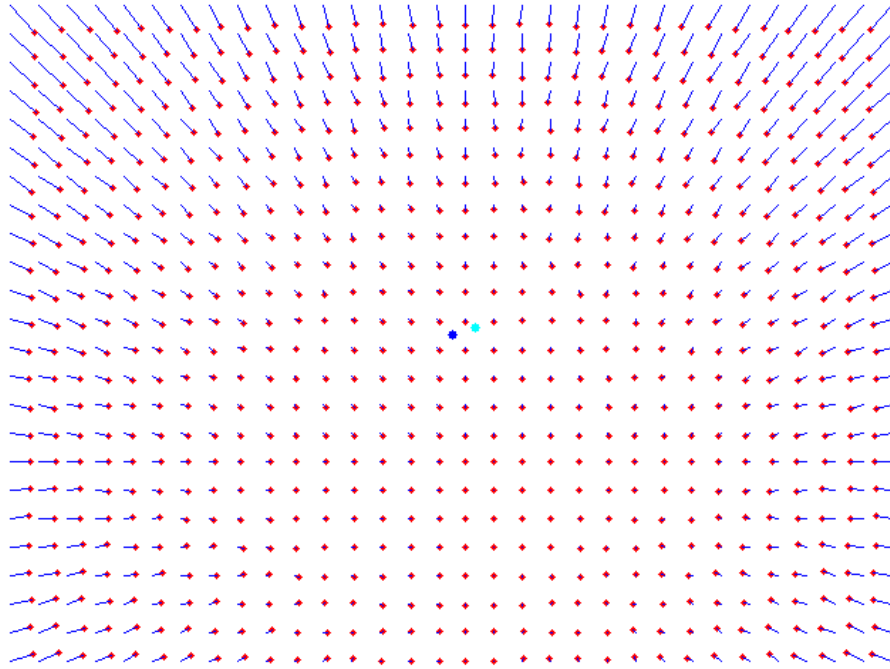


Bild 3.17: Rektifizierungsmatrix für die rechte Kamera

Die roten Punkte beschreiben die Position eines Pixels. Die blaue Linie bezeichnet den Vektor um den das Pixel verschoben werden muss. Zusätzlich ist das Zentrum des Adressbereichs als blauer Punkt und der Hauptpunkt als cyangefärbter Punkt eingezeichnet. Die Abbildung 3.18 zeigt ein mit dem eingesetzten Kamerasystem aufgenommenes Bildpaar vor und nach der Rektifizierung.

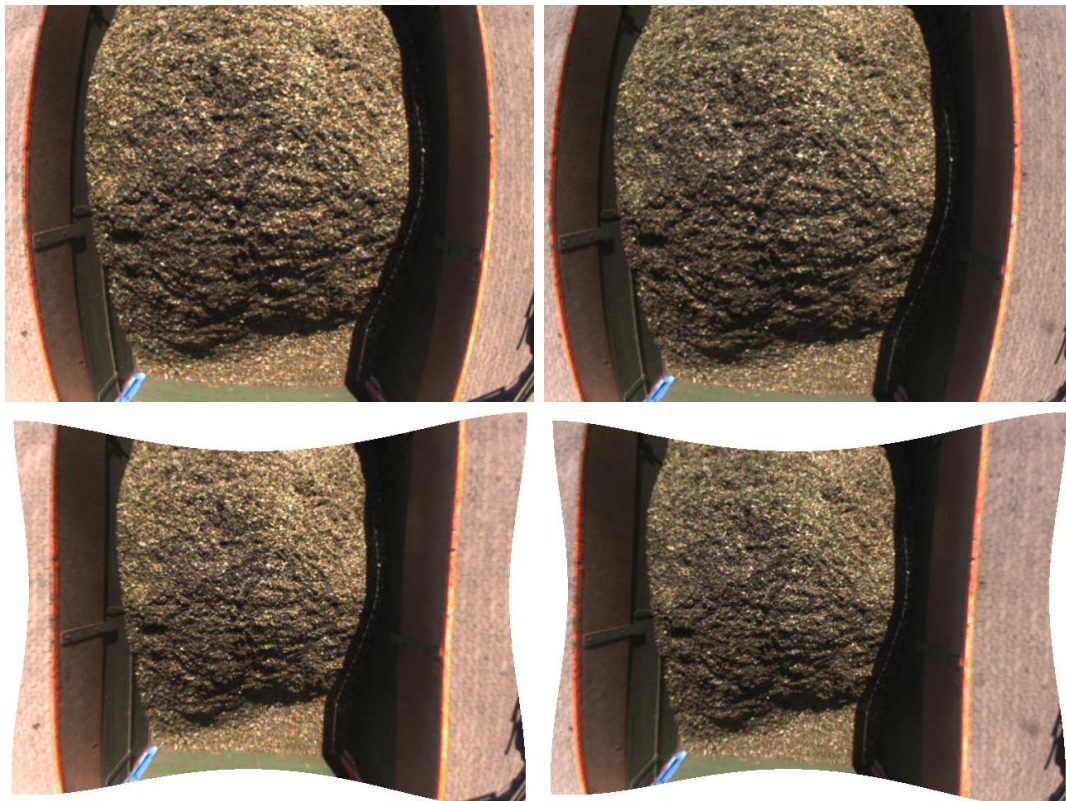


Bild 3.18: Beispiel einer Stereoansicht und ihrer Rektifikation

An den rektifizierten Bildern wird deutlich, dass der Einfluss der Linsenverzeichnung erheblich reduziert wurde. Während in den nicht rektifizierten Bildern (oben) die senkrechten Linien besonders an den Rändern eine deutliche Krümmung aufweisen, erscheinen die gleichen Linien in den rektifizierten Bildern gradliniger. Es ist naheliegend, dass anhand der rektifizierten Bilder die automatische Erkennung von geometrischen Formen einfacher ist. Darüber hinaus wird deutlich, dass die korrespondierenden Punkte der beiden Bilder im rektifizierten Bildpaar auf einer horizontalen Linie liegen, während in den Originalbildern ein Versatz entlang der Y -Achse offensichtlich ist. Die Bedeutung der Rektifizierung für diesen Prozess wird in den folgenden Kapiteln 4 und 5 über die Segmentierung sowie die stereoskopische Auswertung der Bilddaten eingehend erläutert.

4 Segmentierung

Die Segmentierung von Bilddaten ist ein grundlegendes Problem für die meisten Anwendungen in der digitalen Bildverarbeitung [32, 33, 34, 35]. Als Segmentierung kann man die Partitionierung von Bilddaten in Bereiche, die ein gleiches oder ähnliches Verhalten zeigen, bezeichnen. Im Allgemeinen wird die Segmentierung zur Extraktion der interessierenden Bildbereiche oder für die Zuordnung von Bildbereichen zu Objekten angewendet. Die meisten Algorithmen zur Bildsegmentierung basieren auf den zwei grundlegenden Eigenschaften der Ähnlichkeit und Änderung des betrachteten Merkmals.

Die Methoden, welche die Änderung des betrachteten Merkmals nutzen, werden als *boundary-based*-Methoden [33] bezeichnet. Diese Methoden verwenden die Information des Gradienten der Merkmalsintensität, um den Übergang von einem Bildmerkmal in ein anderes Bildmerkmal zu erkennen. So kann durch die Detektion der Grenze, des Bereiches, welches das interessierende Bildmerkmal zeigt, dieser indirekt identifiziert werden.

Die Methoden, welche die Ähnlichkeit eines Bildbereiches bezüglich des betrachteten Merkmals nutzen, werden als *region-based*-Methoden [32] bezeichnet. Diese Methoden gruppieren die Pixel, die hinsichtlich des betrachteten Merkmals in definierten Toleranzen ähnlich sind. Häufig werden Segmentierungsverfahren, die diese Methode nutzen, auf die Grau- oder Farbwerte eines Pixels angewendet. Für den hier vorliegenden Fall, sowie jede Form der Klassifizierung und Segmentierung von natürlichen Oberflächen, werden die sichtbaren Texturen als Merkmal verwendet.

Mit dem Ziel, die zu verarbeitende Datenmenge frühzeitig zu reduzieren, sollen die interessierenden Bildbereiche (*Regions of Interest ROI*) aus den Bildern extrahiert werden. Dies betrifft einerseits die Ränder des Laderaumes und andererseits die Bildbereiche, die das Ladegut zeigen. Der interessierende Bildbereich der Bilder ist verschachtelt. In einem ersten Schritt werden die Ränder des Laderaumes extrahiert. Mit dieser Information werden die Bildbereiche unterdrückt, die den Hintergrund zeigen. Aus dem verbleibenden Bild werden dann die Bildbereiche, die das Ladegut zeigen, ausgeschnitten.

4.1 Der Prozess der Segmentierung

In einem ersten Schritt wird mit Hinsicht auf die zu verarbeitende Datenmenge eine Reduzierung der Bilddaten auf die interessierenden Bildbereiche angestrebt. Dies sind einerseits die Ränder des Laderaumes und andererseits das Ladegut selbst. Die Identifikation der Bildbereiche, welche die Ränder zeigen, wird durch eine Farbfilterung erreicht. Anschließend werden aus dem Bild, das sich aus der Farbfilterung ergibt, die Kanten der identifizierten Ränder mit Hilfe einer *CANNY-Kanten-Detektion* [9] identifiziert.

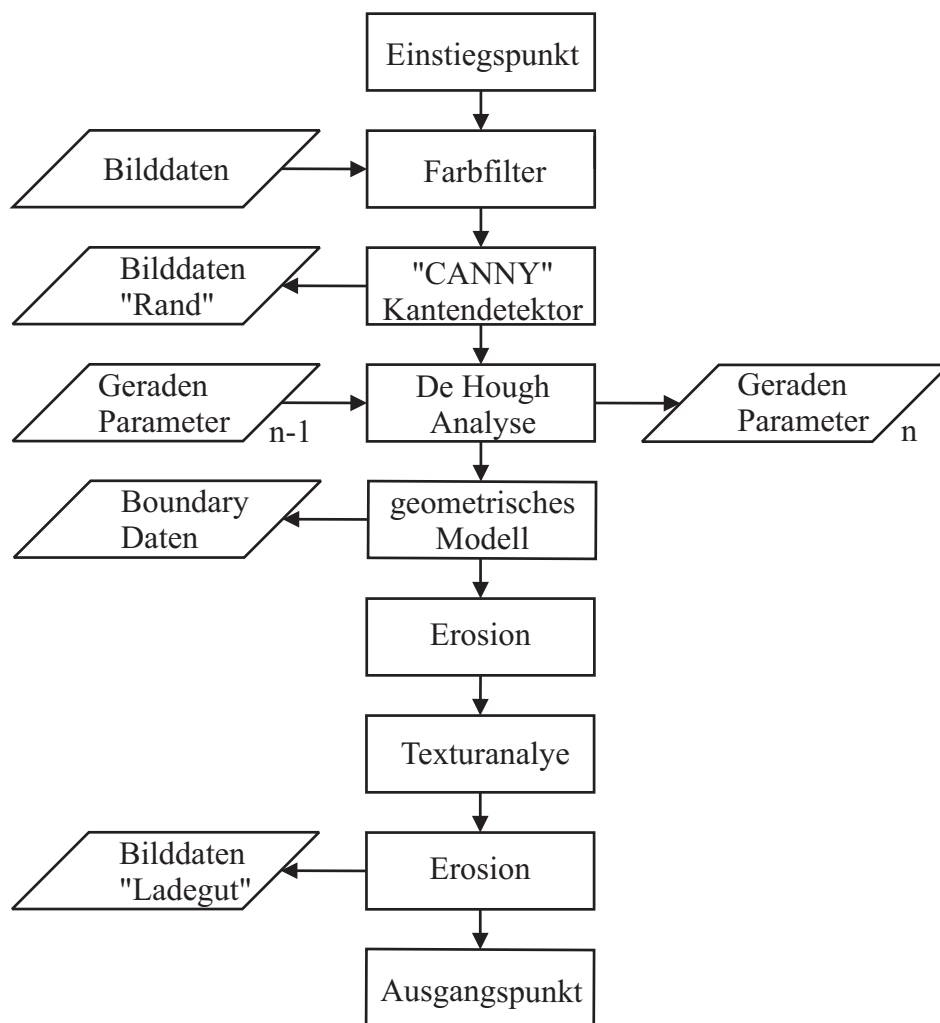


Bild 4.1: Flussdiagramm des Funktionsblocks Segmentierung

Das Ergebnis der Kantendetektion ist ein Binärbild, dessen markierte Pixel die Position der Ränder repräsentieren.

Aus diesem Binärbild werden nun die logischen Indizes der markierten Bildpunkte an die DeHough-Analyse übergeben. Die DeHough-Analyse wiederum untersucht die Eingabedaten auf die Existenz von Geraden. Hierfür werden dem DeHough-Algorithmus die Parameter der im vorherigen Prozessdurchlauf gefundenen Geraden bereitgestellt. Eine Beschreibung der Details des implementierten DeHough-Algorithmus erfolgt in Kapitel 4.2.3. Die so gefundenen Geraden werden dann durch die Anwendung des Modells eines allgemeinen Vierecks in einen geschlossenen Polygonzug umgewandelt. Es findet so eine Reduzierung der vorliegenden Geraden auf die Strecken zwischen den Eckpunkten des Polygons statt. Abschließend wird die vom Polygonzug eingeschlossene Fläche ermittelt. Mit dieser Information wird durch die Erosion der Bildbereich identifiziert, welcher das Innere des Laderaumes zeigt. Darüber hinaus liefert die Anwendung des geometrischen Modells den minimal benötigten Adressbereich des Bildes. Mit dieser Information wird der einzuziehende Bildbereich für das nächste Bild begrenzt. Die Details des geometrischen Modells werden in Abschnitt 4.2.6 erläutert. Aus dem verbleibenden Bild wird nun durch eine Analyse der lokalen Texturen der Bildbereich, der das Ladegut zeigt, identifiziert und durch Erosion extrahiert. Die Einzelheiten der Texturanalyse werden in Kapitel 4.3.1 beschrieben. Die Abbildung 4.1 zeigt das Flussdiagramm der Segmentierungsfunktion sowie deren relevante Ein- und Ausgabedaten.

4.2 Interessierende Bildbereiche

Für die Bemessung des Füllstandes soll ein dreidimensionales Modell der Ladegutoberfläche generiert werden. Da sich das Kamerasystem relativ zum Laderaum bewegt, muss für den Laderaum und somit auch für das Ladegut ein ortsfestes Koordinatensystem definiert werden. Die Daten des Schüttgutmodells sollen dann in das Koordinatensystem des Laderaumes transformiert werden (vgl.6). Das Koordinatensystem des Laderaumes wird sinnvollerweise durch den Rand des Laderaumes definiert. Es ist also notwendig, die Pixel, welche die Ränder des Laderaumes zeigen, zu identifizieren und dem Objekt *Rand* zuzuordnen. Für die Extraktion der Bildbereiche, die das Ladegut zeigen, ist es sinnvoll, zunächst den Bildbereich, der die Innenseite des Laderaumes zeigt, zu identifizieren, um aus diesen Daten dann das Ladegut zu extrahieren. Diese Pixel werden dem Objekt *Ladegut* zugeordnet. Die Abbildung 4.2 verdeutlicht die Segmentierung und Objektzuordnung der interessierenden Bildbereiche.

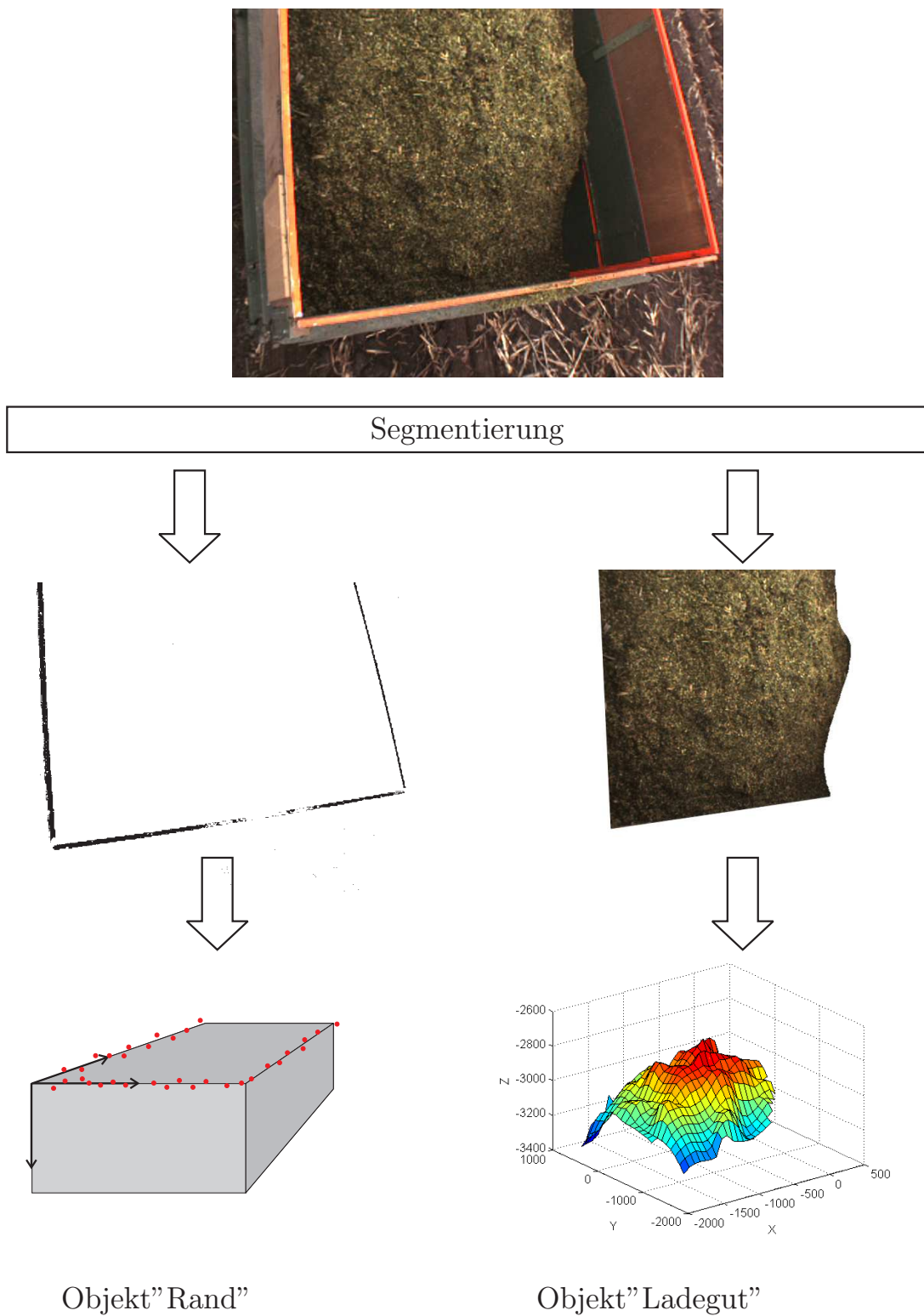


Bild 4.2: Segmentierung der interessierenden Bildbereiche

4.2.1 Identifikation der Ränder

Die Information über die Lage der Ränder des Laderaumes ist von entscheidender Bedeutung für den Messprozess. Es ist naheliegend, dass mit der Kenntniss der Ränder des Laderaumes, der für die weitere Bildverarbeitung interessierende Bildbereich definiert werden kann. Darüber hinaus kann anhand der Ränder das Koordinatensystem des Laderaumes identifiziert werden. Da sich das Messobjekt während der Messung relativ zum Kamerasystem bewegt, ist die Kenntniss des Koordinatensystems des Laderaumes von grundlegender Bedeutung für die Bemessung der Füllhöhe. Eine detaillierte Beschreibung der Definition des Koordinatensystems des Laderaumes, sowie der Bemessung der lokalen Füllhöhe, ist in Kapitel 6 aufgeführt. Um die Bildbereiche, die das Innere des Laderaumes zeigen, identifizieren zu können, werden in einem ersten Schritt die Ränder des Laderaumes über die in Abschnitt 4.2.2 beschriebene Methode der Farbfilterung identifiziert. Hierfür wurden zuvor die Ränder des Laderaumes rot markiert. Das Ergebniss der Farbfilterung ist ein Binärbild, in dem alle TRUE gesetzten Pixel die Bildbereiche, die den Rand zeigen, repräsentieren. Die Abbildung 4.3 zeigt das Ergebnis der Farbfilterung.

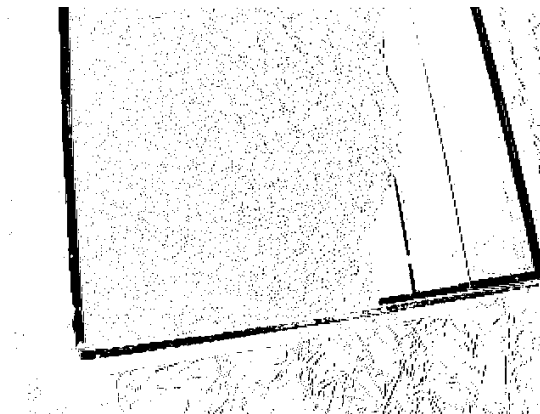


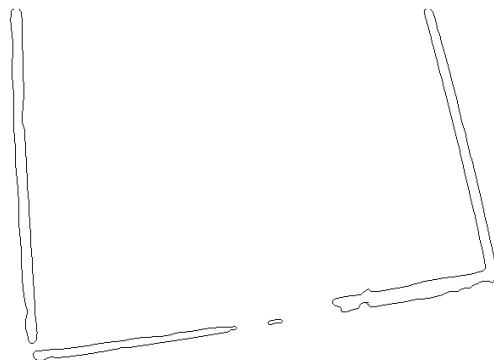
Bild 4.3: Ergebnis der Farbfilterung

Es sind noch erhebliche Störungen sichtbar, die aber durch eine zweidimensionale Medianfilterung unterdrückt werden. In Versuchen hat sich eine Medianfilterung mit einem 10×10 -Punkte-Kernel als geeignet gezeigt. Die Abbildung 4.4 zeigt das Ergebnis der Farbfilterung mit anschließender Medianfilterung. Aus der Abbildung 4.3 wird deutlich, dass die schwarzmarkierten Bereiche, die den Rand repräsentieren, Linien mit einer Stärke von mehr als einem Pixel darstellen. Um den Verlauf der abgebildeten Ränder zu definieren, ist es aber hinreichend, eine Linie mit einer Stärke vom einem Pixel zu erzeugen.



Bild 4.4: Ergebnis der Median-Filterung

Alle zusätzlich als Rand identifizierten Pixel tragen nicht zur Definition der dargestellten Geraden bei. Weitere markierte Pixel hätten lediglich eine unnötige Vergrößerung der weiter zu verarbeitenden Datenmenge zur Folge. Um nun die Menge der markierten Pixel, die der DeHough-Analyse übergeben werden, zu reduzieren, ohne den Informationsgehalt bezüglich der dargestellten Ränder zu schmälern, wird auf das Ergebnis der Farbfilterung ein *CANNY*-Kantendetektor [9] angewendet. Der Kantendetektor verstärkt durch Überlagerung der ersten und zweiten Ableitung der Bilddaten die Bildbereiche, an denen Intensitätsänderungen bzw. Farbänderungen auftreten. Somit werden die Bildbereiche verstärkt, die wiederum die Kanten der Ränder des Laderaumes repräsentieren. Die Abbildung 4.5 zeigt das Bild nach der Anwendung des *CANNY*-Kantendetektors.

Bild 4.5: Ergebniss der *Canny* Kantendetektion

Auch in diesem Ergebnis (Abbildung 4.5) verbleiben markierte Pixel, die redundante Informationen bezüglich der Lage der Ränder enthalten. Mit Hinsicht auf die Berechnungszeit wird die zu verarbeitende Datenmenge aber wie in den Abbildungen 4.3 und 4.5 deutlich reduziert. Eine weitere Reduzierung der markierten Pixel wäre bezüglich des Mehraufwandes für die Berechnung der redundanten Pixel mit unverhältnismäßig großem Aufwand verbunden.

4.2.2 Farberkennung bei variabler Beleuchtung

FILLED soll unter natürlich variierenden Beleuchtungsverhältnissen betrieben werden können. Das Sonnenlicht kann im Bereich des sichtbaren Lichtes als annähernd ideal polychromatisch angesehen werden. Unter dieser Voraussetzung kann davon ausgegangen werden, dass die Folge variierender Beleuchtung eine entsprechende Änderung der Luminanz ist. Um nun unabhängig von der Luminanz des zu filternden Farbtönen stabil eine Farbe erkennen zu können, ist es sinnvoll, die Verhältnisse der Intensitäten der beteiligten Farbkomponenten eines Pixels auszuwerten. Bezeichnet man den Ort eines Pixels $P_{u,v}$ in einem Bild mit u, v , so kann die Intensität des betrachteten Pixels mit $I(u, v)$ bezeichnet werden. In den beteiligten Farbebenen des digitalisierten Bildes werden die Intensitäten eines Pixels mit $I_R(u, v)$ für den Rot-Layer, $I_G(u, v)$ für den Grün-Layer und $I_B(u, v)$ für den Blau-Layer bezeichnet. Um die roten Ränder des Laderaumes aus den Bildern zu extrahieren, werden die Quotienten aus Rot- und Grün-Layer sowie Rot- und Blau-Layer ausgewertet. Der Wert $I(u, v)$ eines Pixels des gefilterten Bildes ergibt sich nach:

$$I(u, v)_{rim} = \begin{cases} 1 & q_G(u, v) \geq T_G \wedge q_B(u, v) \geq T_B \\ 0 & \text{sonst} \end{cases} \quad (4.1)$$

$$q_G(u, v) = \frac{I_R(u, v)}{I_G(u, v)} \quad (4.2)$$

$$q_B(u, v) = \frac{I_R(u, v)}{I_B(u, v)} \quad (4.3)$$

Da sich im allgemeinen Fall bei einer deutlichen Änderung der Beleuchtungssituation auch die Farbtöne ändern, wird über einen Regler der Durchtrittsfarbtönen der Beleuchtungssituation nachgeführt. Hierfür wird in regelmäßigen Abständen der Mittelwert der Farbquotienten q_G und q_B aller Pixel der gefilterten Bereiche gebildet und die in Gleichung 4.1 beschriebenen Parameter T_G und T_B der Filtervorschrift angepasst.

4.2.3 Die DeHough-Transformation

Die DeHough-Analyse ist ein Werkzeug, das es ermöglicht, in einer Menge von ausgewählten Pixeln geometrische Formen zu erkennen. Im vorliegenden Fall wird die DeHough-Analyse verwendet, um die Kanten des Laderaumes in Geraden die in Polardarstellung angegeben werden, zu transformieren. Im Rahmen der DeHough-Analyse wird geprüft, ob die durch vorhergehende Filterung markierten Pixel Elemente einer Geraden sind. Zunächst müssen die Koordinaten der ausgewählten n -Pixel ermittelt werden. Diese werden in dem $2 \times n$ Vektor \vec{P}_m abgelegt. Nun sieht die DeHough-Transformation vor, zu jedem Punkt P_{mi} aus \vec{P}_m eine beliebig große Menge Geraden $G_{P_{mi}}$ durch diesen Punkt zu definieren. Jede dieser Geraden wird durch die Parameter ρ_{mij} und φ_{mij} in Polardarstellung beschrieben. Der Satz der Geradenparameter durch einen Punkt P_{mi} wird als Akkumulatorzelle A_{mi} bezeichnet. Liegen nun mehrere Punkte aus \vec{P}_m auf einer Geraden g^* , so weisen diese ein gemeinsames Parameterpaar ρ^* und φ^* auf, welches die Gerade beschreibt, deren Element sie sind. Durch eine Akkumulation gleicher ρ - und φ -Paare über alle Akkumulatorzellen A_{mi} wird der DeHough-Raum \mathbf{D}_i für ein Bild I erzeugt. Die Matrix \mathbf{D}_I entspricht dem zweidimensionalen Histogramm der ρ - und φ -Paare und enthält somit die Häufigkeiten $I_{\rho\varphi}$ aller auftretenden ρ - und φ -Kombinationen. Jedes Element der Matrix \mathbf{D}_I kann einem Parameterpaar ρ und φ und so einer Geraden zugeordnet werden. Der Wert der betrachteten Elemente gibt die Menge der durch diese Gerade akkumulierten Punkte an und ist somit ein Maß für die Bestimmtheit der Lösung.

4.2.4 Berechnung des DeHough- Raumes

Es ist sinnvoll, die Menge der Geraden $G_{P_{mi}}$ um ein markiertes Pixel P_{mi} in diskreten äquidistanten Winkelschritten $\delta\varphi = \text{const}$ zu bilden. Die Abbildung 4.6 zeigt die Konstruktion einer Geraden durch einen Punkt in Polardarstellung. Der Abstand $\rho(P_{mi}, \varphi)$ einer Geraden $g_{P_{mi}}$ durch den Punkte P_{mi} zum Ursprung des Koordinatensystems $\vec{0}$ ergibt sich nach:

$$\rho(P_{mi}, \varphi) = \frac{\vec{A}_g \times \vec{c}(\varphi)}{|\vec{c}(\varphi)|} \quad (4.4)$$

Der Vektor \vec{A}_g zeigt vom Ursprung des Koordinatensystems zu einem beliebigen Punkt P_g der Geraden g . Der Vektor $\vec{c}(\varphi)$ definiert die Richtung der Geraden. Als Aufpunktvektor wird hier der Ortsvektor $\vec{R}_{P_{mi}}$ zum betrachteten Pixel P_{mi} verwendet. Der Vektor $\vec{c}(\varphi)$ ist eine Funktion des gewählten Winkels φ .

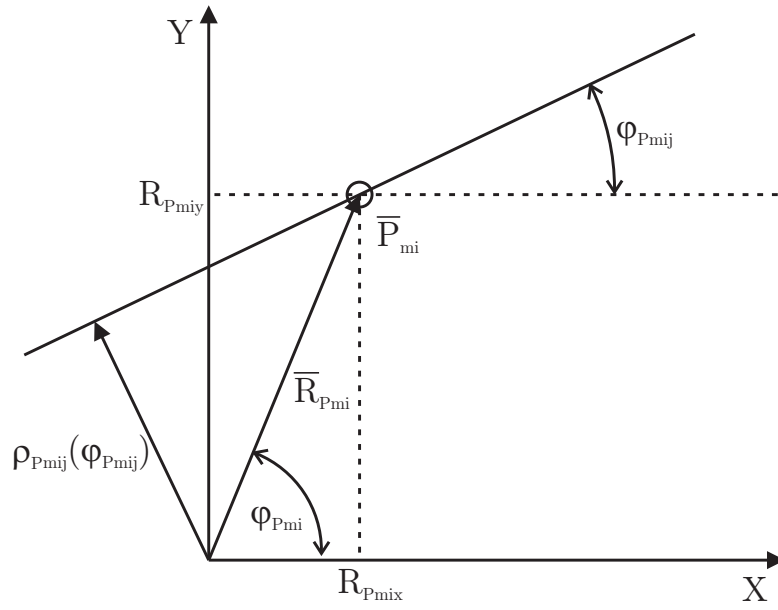


Bild 4.6: Gerade in Polardarstellung

\vec{c} wird als Einheitsvektor gebildet, sodass sich die Länge $\rho(\vec{P}_{mi}, \varphi)$ des lotrechten Aufpunktvektors als Funktion des Winkels φ nach Gleichung 4.5 ergibt.

$$\begin{aligned}
 \rho(P_{mi}, \varphi) &= \vec{R}_{P_{mi}} \times \vec{c}(\varphi) \\
 \rho(P_{mi}, \varphi) &= |\vec{R}_{P_{mi}}| \begin{pmatrix} \cos(\varphi_0) \\ \sin(\varphi_0) \end{pmatrix} \times \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix} \\
 \rho(P_{mi}, \varphi) &= |\vec{R}_{P_{mi}}| \sin(\varphi + \varphi_P) \quad 0 \leq \varphi \leq \pi
 \end{aligned} \tag{4.5}$$

Hierbei ergibt sich φ_P nach

$$\varphi_P = \arctan\left(\frac{R_{P_{mi},y}}{R_{P_{mi},x}}\right) \tag{4.6}$$

Gleichung 4.5 definiert die Menge aller Geraden $G_{P_{mi}}$ um das betrachtete Pixel P_{mi} . Man erhält für jedes markierte Pixel P_{mi} einen Satz von Gleichungen, welche die Menge der Aufpunktvektoren $\vec{R}_{P_{mi}}$ als Funktion der n_{max} Winkel Φ definiert. Die Akkumulatorzelle $A_{P_{mi}}$ zu einem Punkt P_{mi} ergibt sich nun durch die Anwendung der Gleichung 4.5 auf den Vektor $\vec{\Phi}$ der diskreten Winkel.

$$\vec{\Phi}_n = n \frac{\pi}{n_{max}} \quad n = 0, 1, 2, \dots, n_{max} - 1 \tag{4.7}$$

$$A_{P_{mi}} = |\vec{R}_P| \sin(\vec{\Phi} + \varphi_{P_{mi}}) \tag{4.8}$$

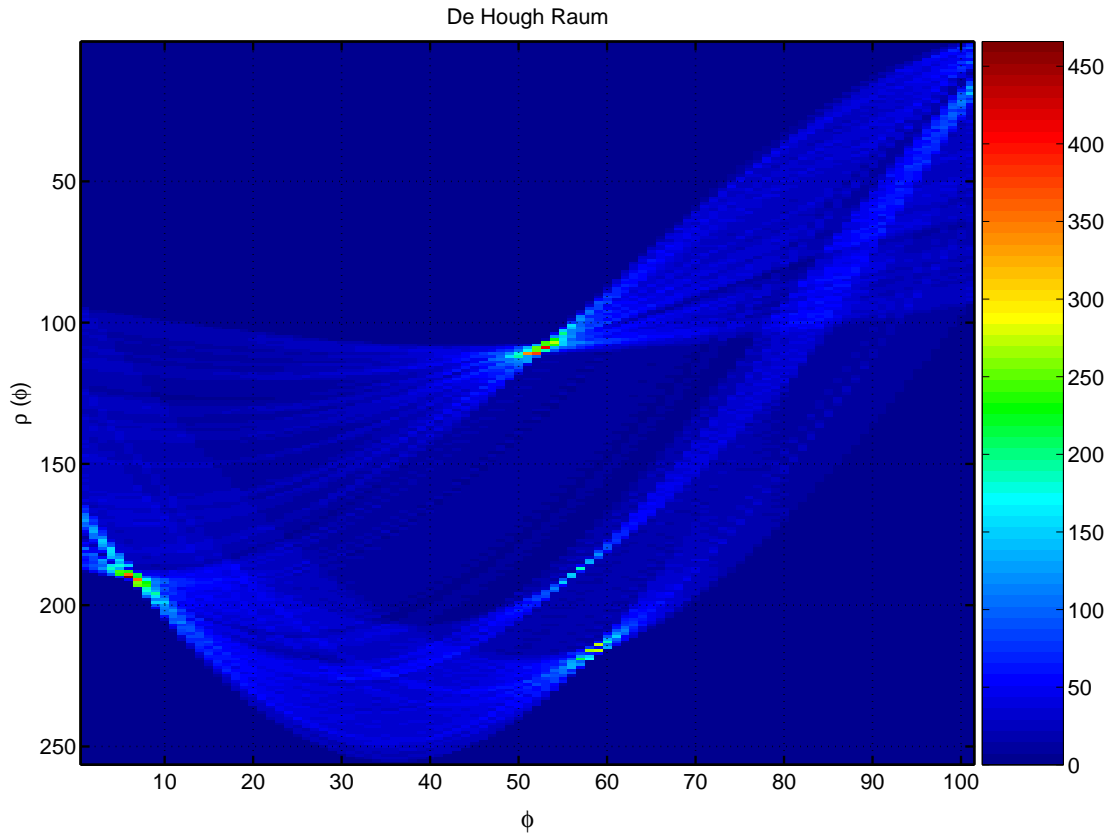


Bild 4.7: DeHough Raum

Der Wert n_{max} gibt die Anzahl der äquidistanten Diskretisierungsschritte der geraden Winkel an. Für die Erfassung aller möglichen Geraden ist es hinreichend, einen Winkelbereich von φ_0 bis $\varphi_{max} = \varphi_0 + \pi$ zu diskretisieren. Alle weiteren Winkel würden redundante Ergebnisse liefern, die sich lediglich im Vorzeichen unterscheiden. Dieses Verfahren liefert für jedes markierte Pixel eine Akkumulatorzelle. Durch eine Histogrammauswertung aller Akkumulatorzellen über alle ρ - φ -Kombinationen erhält man den DeHough-Raum. Die Abbildung 4.7 zeigt einen DeHough-Raum für ein Bild, in dem vier Geraden sichtbar sind. Eine weitere Methode den DeHough-Raum zu berechnen, nutzt die Möglichkeit die exakte Geradengleichung, die durch zwei Punkte definiert wird, zu berechnen und anschließend im Rahmen der Histogrammerstellung eine Diskretisierung des Winkels φ und der Länge des Aufpunktvektors ρ zu berechnen. Dieses Verfahren bietet eine höhere Genauigkeit für die Bestimmung der Geradenparameter, da die Geradengleichungen zunächst als Gleitkommawerte berechnet und anschließend diskretisiert werden. Die Parameter einer gemeinsamen Gerade g_{P_i, P_j} zweier Punkte P_i und P_j findet man bekanntlich durch Gleichsetzen der zugehörigen Geradengleichungen.

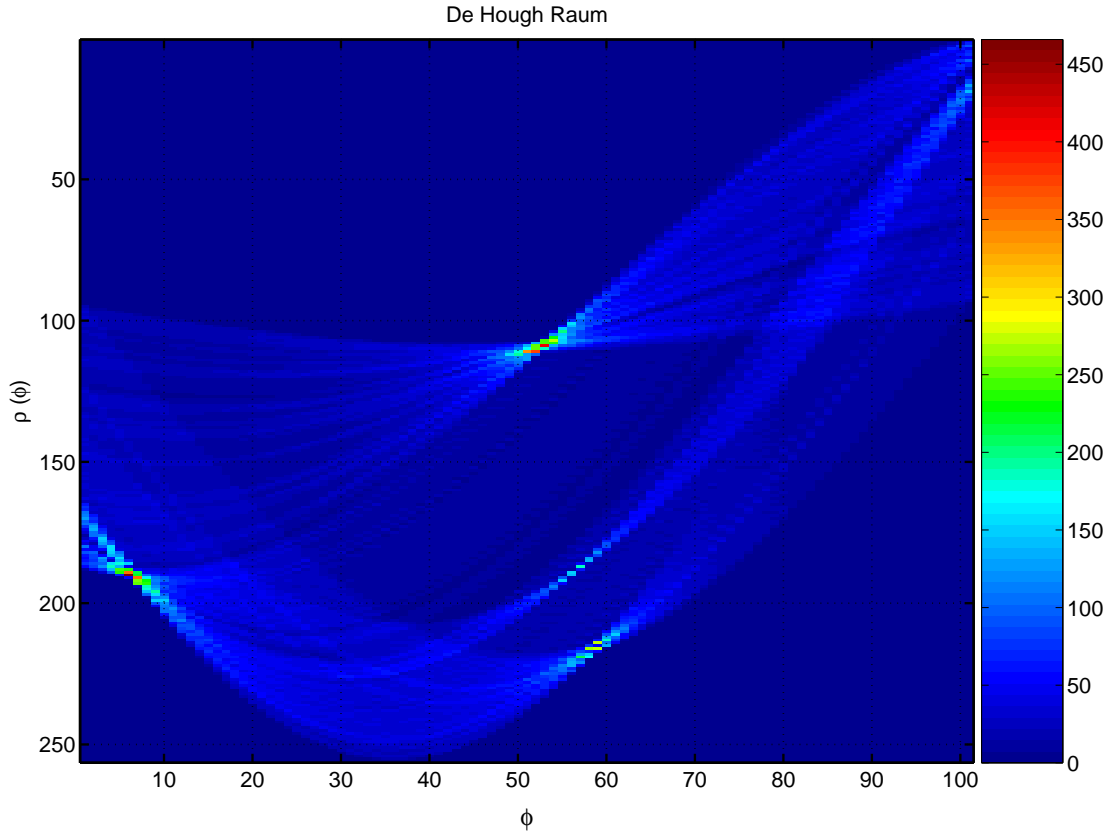


Bild 4.8: Analytisch bestimmter De Hough Raum

Die Gleichung der Geraden g_{P_i, P_j} ergibt sich nach:

$$\rho(P_i, \varphi_i) = \rho(P_j, \varphi_j) \quad (4.9)$$

$$\varphi_i = \varphi_j \quad (4.10)$$

$$\varphi_{i,j} = \arctan \left(\frac{|R_{P_i}| \cos(\varphi_i) - |R_{P_j}| \cos(\varphi_j)}{|R_{P_i}| \sin(\varphi_i) - |R_{P_j}| \sin(\varphi_j)} \right) \quad (4.11)$$

$$\rho_{ij} = |R_{P_i}| \sin(\varphi_{ij} + \varphi_i) \quad (4.12)$$

Durch die Bildung aller möglichen Punktpaare erhält man die Menge aller durch die markierten Pixel definierten Geraden. Nun kann durch eine Histogrammauswertung über alle ρ - φ -Kombinationen eine Aussage über die Existenz von sichtbaren Linien im Bild getroffen werden. Die Abbildung 4.8 zeigt den DeHough-Raum. Hier wurde der Winkel der Geraden $\varphi = 0.. \pi$ auf 100 Werte und der Wertebereich für die Länge des Aufpunktvektors ρ auf 256 Werte abgebildet. Nachteilig hierbei ist, dass die Menge der zu verarbeitenden Daten mit

dem Quadrat der markierten Pixel wächst. Die Anzahl der zu bildenden Punktepaare n_P ergibt sich als Funktion der Anzahl der markierten Pixel m nach:

$$n_P = \frac{m^2 - m}{2} \quad (4.13)$$

Bei der *nicht analytischen* Methode, ergibt sich die Anzahl der zu verarbeitenden Werte durch Multiplikation der Anzahl der markierten Pixel mit der Anzahl der Diskretisierungsschritte.

$$n = mn_{dmax} \quad (4.14)$$

Da die Berechnung des Eintrags einer Akkumulatorzelle annähernd identisch mit der Berechnung einer Geradengleichung ist, kann die Anzahl der zu verarbeitenden Werte als sinnvolles Maß für den Vergleich der Berechnungsdauer der beiden Verfahren genutzt werden. Mit Hinsicht auf die begrenzte Zykluszeit wird hier die erstgenannte Methode eingesetzt. Diese liefert zwar eine geringere Genauigkeit, sie benötigt dafür eine deutlich geringere Berechnungszeit. Wie bereits am Anfang dieses Abschnitts erwähnt, entspricht ein Wert des DeHough Raumes an einer Stelle ρ_o, φ_o der Menge der Pixel, die Element einer Geraden mit diesen Parametern sind. Die Länge einer Linie im Bild bestimmt also maßgeblich die Ausprägung des zugehörigen Maximums des DeHough Raumes. Die Entscheidung ob eine Linie erkannt wird, wird auf Basis der Ausprägung des gefundenen Maximums des De-Hough-Raumes getroffen. Unterschreitet die Zahl C_{gp} der auf einer Linie g akkumulierten Punkte einen kritischen Wert C_{gpkrit} , so wird diese Lösung verworfen.

4.2.5 Optimierung der DeHough-Transformation

Das Messsystem *FILLED* soll den Füllstand während der Beladung messen. Es ist das Ziel, rechtzeitig in den Beladungsprozess eingreifen zu können. Eine wesentliche Anforderung an das System ist es also, die Dauer für die Berechnung einer Lösung zu begrenzen. Durch die Optimierung der DeHough-Transformation soll, bei gleichbleibender Qualität der Messwerte, die Dauer des Messvorgangs zu reduziert werden. Dieses kann über eine Vorsteuerung des Winkelbereichs der Akkumulatorzellen und eine rechtzeitige Klassifikation nach signifikanten und nicht signifikanten Pixeln erreicht werden. Im Abschnitt 4.2.6 wird beschrieben, wie mit Hilfe eines geometrischen Modells der Bildbereich identifiziert wird, der das Innere des Transportmittels zeigt. Die Bildbereiche, die nicht das Innere des Laderaumes einschließlich seiner rot markierten Kanten zeigen, brauchen nicht weiter betrachtet zu werden. Unter der

Annahme, dass sich die relative Position des Laderaumes zum Kamerakoordinatensystem in einer Zykluszeit nur geringfügig ändert, ist es möglich, vom momentan interessierenden Bildbereich auf den interessierenden Bildbereich des folgenden Bildes zu schließen. Abbildung 4.9 zeigt die Definition des geschätzten interessierenden Bildbereichs für das folgende Bild.

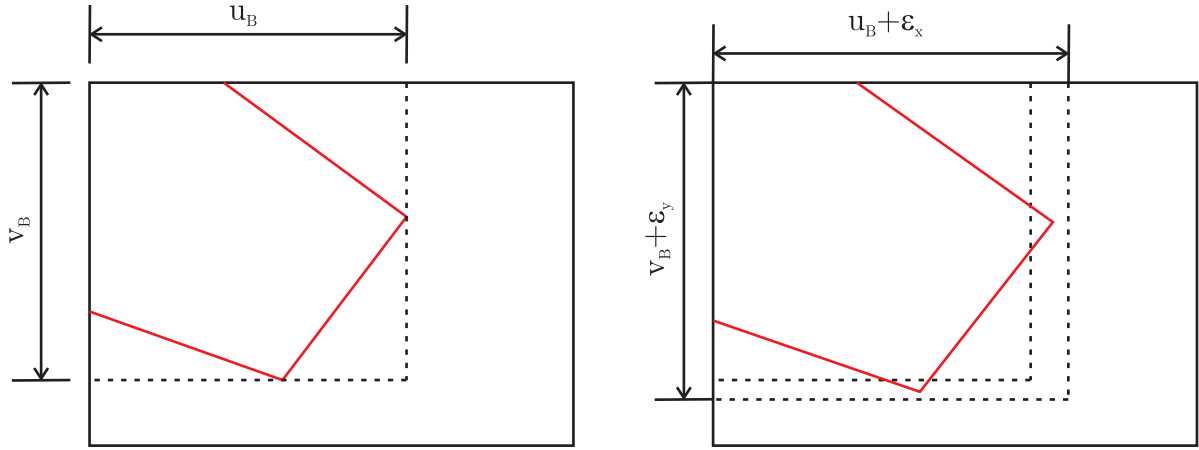


Bild 4.9: Begrenzung des Bildbereichs(links), ROI im neu eingezogenen Bild

Die Koordinaten des angenäherten interessierenden Bildbereichs ergeben sich nach:

$$u'_{B,max} = u_{B,max} + \epsilon_u \quad (4.15)$$

$$v'_{B,max} = v_{B,max} + \epsilon_v \quad (4.16)$$

$$u'_{B,min} = u_{B,min} - \epsilon_u \quad (4.17)$$

$$v'_{B,min} = v_{B,min} - \epsilon_v \quad (4.18)$$

mit

$$u_{B,max} = \max_{i \rightarrow 4}(\mathbf{S}_{u,i}) \quad (4.19)$$

$$v_{B,max} = \max_{i \rightarrow 4}(\mathbf{S}_{v,i}) \quad (4.20)$$

$$u_{B,min} = \min_{i \rightarrow 4}(\mathbf{S}_{u,i}) \quad (4.21)$$

$$v_{B,min} = \min_{i \rightarrow 4}(\mathbf{S}_{v,i}) \quad (4.22)$$

Hierbei bezeichnet \mathbf{S} die 2×4 Matrix der 4 Schnittpunkte der gefundenen Geraden(vgl. Kapitel 4.2.6). S_u und S_v bezeichnen demnach die Spaltenvektoren, welche die u - bzw. v -Komponenten der gefundenen Schnittpunkte enthalten. Auf diese Weise können alle Punkte, die aufgrund von Rauschen in den nicht interessierenden Bildbereichen durch den Farbfilter

markiert werden, von der Analyse der DeHough Transformation ausgeschlossen werden. Diese Reduzierung der markierten Bildpunkte beeinflusst die Berechnungsdauer der DeHough-Transformation erheblich, da für jeden Punkt eine Akkumulatorzelle berechnet werden muss. Eine weitere Methode die Datenmenge sinnvoll zu reduzieren, bietet die Vorsteuerung des Abtastbereichs der Geradenwinkel. Hier kann die Bildauswertung sowohl in einer Sequenz, als auch in einem Stereobilddatensatz vorgesteuert werden. Die Abbildung 4.10 zeigt das Prinzip der Vorsteuerung.

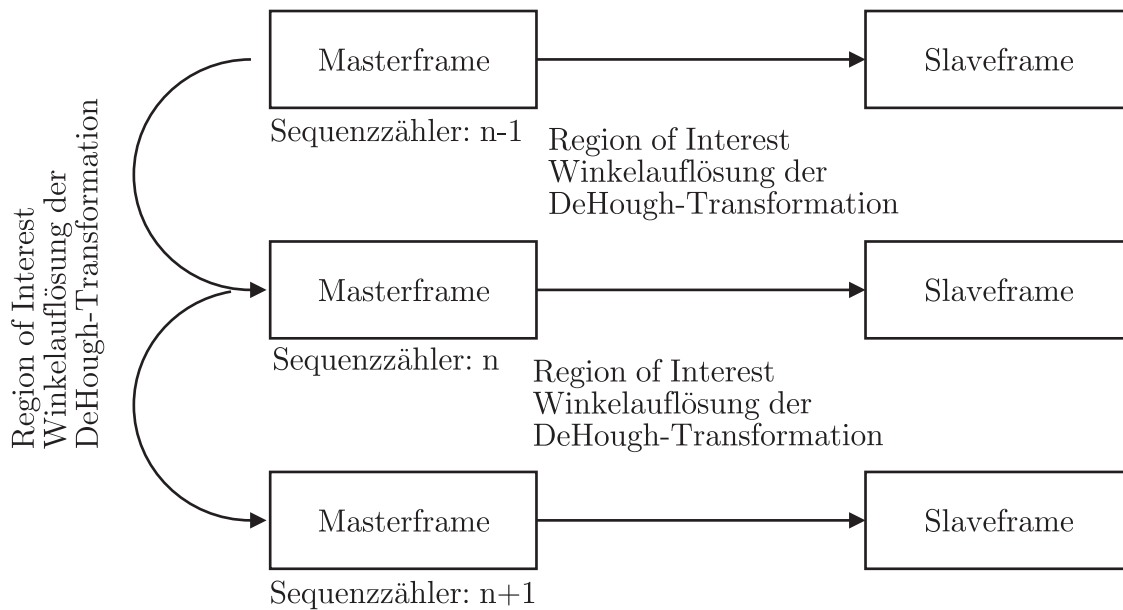


Bild 4.10: Vorsteuerung der Scanbereiche

Zunächst wird die Vorsteuerung des Geradenwinkels für eine Bilderfolge erläutert. Auch hier wird die Annahme getroffen, dass sich die Position und Lage der abgebildeten Geraden im Bild von einem Datensatz zum Folgenden nur wenig ändert. Dies ermöglicht es, den Suchbereich der DeHough Transformation auf eine begrenzte Umgebung um die im vorhergehenden Bild gefundenen Linien einzuschränken. Analog dazu kann man von der Lage der Geraden in einem Bild des Stereopaars auf die Lage der Geraden im anderen Bild schließen. Bedingt durch die Anordnung des Kamerasystems werden die Objekte im Allgemeinen stets mit einem konstanten Versatz in den beteiligten Kameras abgebildet. Dieser Versatz, die Disparität, ist abhängig von der Entfernung der Objekte vom Kamerasystem. Da sich die Entfernung der Ränder zum Kamerasystem während der Überladungsprozesses nur unwesentlich ändert, ist es zulässig, diese Disparität in einer Initialisierungsroutine festzustellen und dann für die weitere Messung zu verwenden. Die Vorsteuerung bietet zwei Möglichkeiten der Optimierung

des Transformationsprozesses. Durch eine Verkleinerung des zu untersuchenden Winkelintervalls bei konstanter Anzahl von Diskretisierungsschritten, kann bei gleichem Berechnungsaufwand, eine größere Genauigkeit erzielt werden. Verwendet man hingegen gleichbleibend große Diskretisierungsintervalle, so reduziert sich die Anzahl der Diskretisierungsschritte für das reduzierte Intervall. Diese Option bietet die Möglichkeit, bei gleichbleibender Genauigkeit den Berechnungsaufwand zu reduzieren. Vor dem Hintergrund, dass die Objekte durch die Kameras nur mit einer endlichen Auflösung abgebildet werden, ist es sinnvoll, eine begrenzte Auflösung des DeHough-Winkels zu wählen. Aus diesem Grund wird hier der Optimierungsmodus gewählt, der einen reduzierten Berechnungsaufwand bietet. Als problematisch hat sich erwiesen, dass der verwendete Filteralgorithmus zur Kantenverstärkung sowohl die Eintrittskante, als auch die Austrittskante, in einem markierten Bereich verstärkt (vgl. Abbildung 4.5). Man erhält für die gefilterten roten Ränder zwei Markierungen, die von der DeHough-Analyse ausgewertet werden. Demnach erkennt die DeHough-Analyse zwei parallele Linien pro Streckenabschnitt des Laderaumrandes. Dies kommt in den Abbildungen 4.7 durch zwei dicht beieinanderliegende Maxima zum Ausdruck. Eine optimale Lösung stellt nun die Gerade dar, deren Parameter dem Mittelwert der beiden gefundenen Geraden entspricht. Diese Annahme ist zulässig, da die Gerade gesucht wird, die den Verlauf einer Ladebordwandkante optimal repräsentiert. Durch Mittelung der gefundenen Geradenparameter ρ_1 und ρ_2 sowie φ_1 und φ_2 für eine Kante erhält man die Gerade, die hinreichend genau mittig auf dem rot markierten Rand liegt. Die Parameter der Geraden g_i , die eine Kante des Laderaumes repräsentieren, ergeben sich nach Gleichung 4.23.

$$g_i := f(\varphi_i, \rho_i) \quad (4.23)$$

$$\varphi_i = \frac{\varphi_{i1} + \varphi_{i2}}{2} \quad (4.24)$$

$$\rho_i = \frac{\rho_{i1} + \rho_{i2}}{2} \quad (4.25)$$

Da die Existenz einer Geraden im Bild anhand der durch sie akkumulierten Punkte festgemacht wird und die Länge einer Geraden wiederum maßgeblich die Anzahl der durch sie akkumulierten Punkte beeinflusst, werden sehr kurze Linien bei Betrachtung der absoluten Anzahl der akkumulierten Punkte unterdrückt. Dies ist ungünstig, wenn der Laderaum den Sichtbereich der Kameras verlässt. Abbildung 4.11 zeigt eine Situation, in der der Laderaum den Sichtbereich der Kameras soweit verlassen hat, dass zwei seiner Kanten nur als sehr kurze Geraden abgebildet werden. Um diesem Effekt entgegenzuwirken, wird die Länge der gefundenen Geraden auf die maximal mögliche Länge einer kollinearen Geraden im interessierenden Bereich dieses Bildes bezogen.

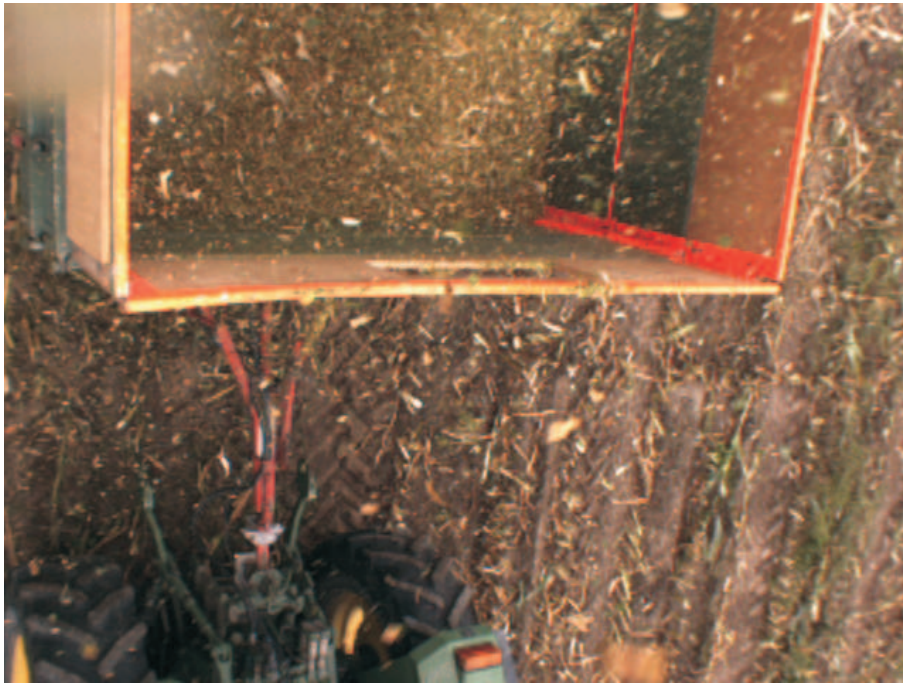


Bild 4.11: Der Laderaum verlässt den Bildbereich

4.2.6 Das geometrische Modell

Die DeHough Transformation ermittelt aus den eingezogenen Bildern die Ränder des Laderaumes als unabhängige Geraden. Mit Hilfe des Modells eines allgemeinen Vierecks sollen nun die gefundenen Geraden auf Strecken, die den sichtbaren Rändern entsprechen, reduziert werden. Darüberhinaus soll so die von diesen Strecken eingeschlossene Fläche identifiziert werden, da diese den interessierenden Bildbereich zeigt. Die meisten Transportmittel haben einen Laderaum mit annähernd rechteckigem Grundriss. Durch die Perspektive der Kameras wird die rechteckige Öffnung des Laderaumes allerdings im Allgemeinen verzerrt abgebildet. Da sich der Blickwinkel des Kamerasystems auf das Transportmittel nur wenig ändert, kann davon ausgegangen werden, dass der Winkelversatz der Geraden zueinander annähernd konstant bleibt. Darüberhinaus kann davon ausgegangen werden, dass wegen der Montage des Kamerasystems die parallelen Ränder des Laderaumes auch annähernd parallel abgebildet werden. Lediglich durch den Einfluss der zentrischen Streckung erscheinen die abgebildeten parallelen Ränder nicht parallel. Diese geometrischen Voraussetzungen ermöglichen eine strukturierte Suche nach den Rändern des Laderaumes. Hierfür wird das globale Maximum des DeHough-Raumes lokalisiert. Wie im Kapitel 4.2.3 bereits erwähnt, entspricht ein Eintrag im DeHough-Raum der Anzahl der Punkte, die Element der Geraden

mit diesen Geradenparametern sind. Das globale Maximum des DeHough-Raumes weist also auf die signifikanteste Linie im Bild hin. Diese wird als Masterlinie identifiziert. Da nach einer bekannten konstanten Geometrie gesucht wird, definiert diese Masterlinie die Lage der gesuchten Geometrie im Bild. Darüber hinaus lässt die Anzahl der Punkte, die Element der gefunden Gerade sind, auf die Länge dieser Geraden schließen. Betrachtet man ein allgemeines Rechteck mit jeweils zwei paarweise parallelen Kanten gleicher Länge, so kann anhand der Masterlinie auf die Ausrichtung des Laderaumes geschlossen werden. Dies ist eine wichtige Information für die Definition des Koordinatensystems des Laderaumes. Die Details hierzu werden in Kapitel 6.4 erläutert. Die Abbildung 4.12 zeigt ein Bild, das an die DeHough Transformation übergeben wird.

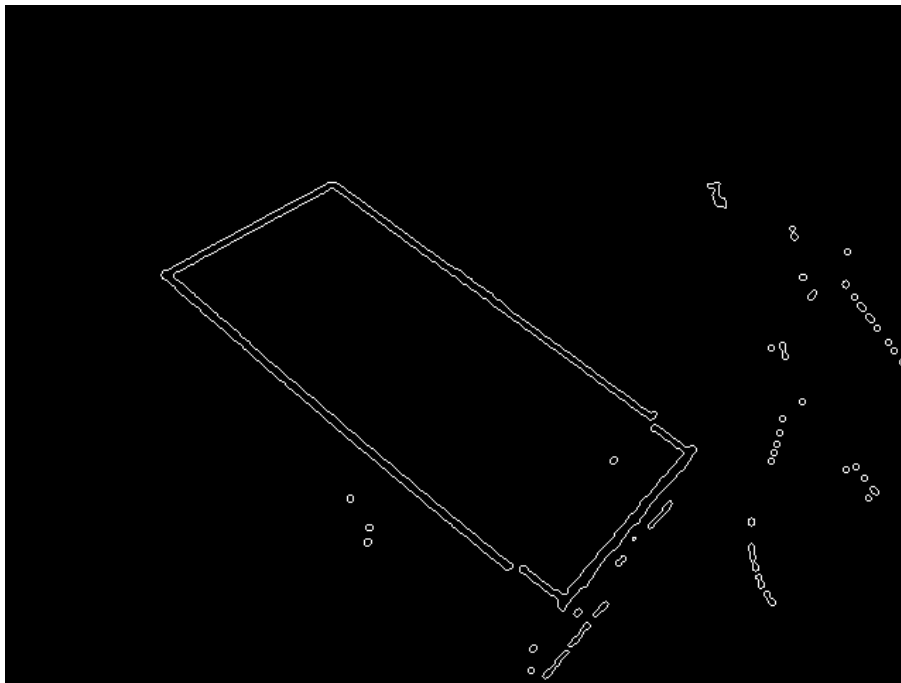


Bild 4.12: Eingabebild in die DeHough-Transformation

Nun wird in der nahen Umgebung um den Winkel φ_{g_M} der Masterlinie eine parallele Linie gesucht, welche den Rand der gegenüberliegenden Ladebordwand identifiziert. Ein weiteres Kriterium für die Existenz des parallelen Randes ist ein hinreichend großer und während des Messvorganges annähernd gleichbleibender Abstand des Maximums entlang der ρ -Achse im DeHough Raum. Im Weiteren wird unter dem Winkel $\varphi_{g_M} + \frac{\pi}{2} \pm \varphi_{offset}$ nach der signifikantesten Linie dieses Winkelbereichs gesucht. Diese schrittweise strukturierte Suche nach definiert

angeordneten Linien im Bild verhindert zusätzlich, dass geradlinige Strukturen im Hintergrund, die nicht zum interessierenden Merkmal des Bildes gehören, als Teile des Laderaumes erkannt werden. Dies betrifft im wesentlichen die üblicherweise reihenförmig angeordneten Maispflanzen sowie Spuren, die vorhergefahrene Fahrzeuge im Boden hinterlassen haben.

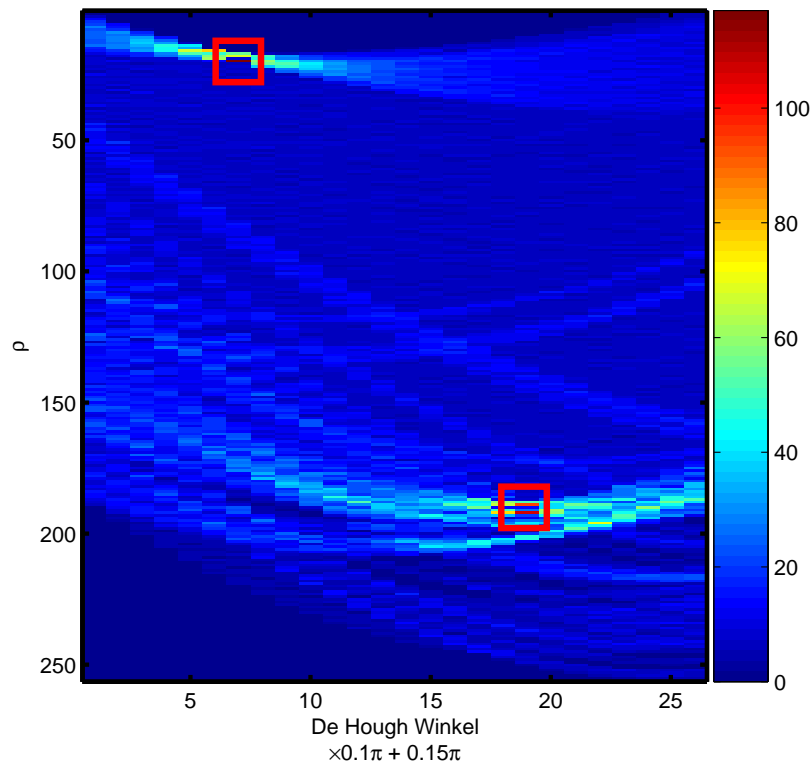


Bild 4.13: DeHough-Raum für einen begrenzten Winkelbereich

Die Abbildung 4.13 zeigt den DeHough Raum für einen begrenzten Winkelbereich. Man erkennt, dass in der Abbildung, dass lediglich zwei Maxima, die den jeweils paarweise parallelen Geraden entsprechen, sichtbar sind.

4.3 Die Extraktion des Schüttgutkegels

Nachdem aus den Bilddaten mit Hilfe der De-Hough Transformation der Bereich, der die Innenseite des Laderaumes zeigt, extrahiert wurde, wird nun der Bildbereich, der das Ladegut zeigt, identifiziert. So wird es möglich, die Bildpunkte und die daraus im Weiteren ermittelten Objektpunkte eindeutig dem Objekt *Ladegut* zuzuweisen. Auf Grund von natürlichen Farbvarianzen und Beleuchtungseinflüssen kann das Ladegut im Allgemeinen nicht anhand seiner Färbung sicher erkannt werden. Es kann jedoch angenommen werden, dass die Partikel des Ladegutes mit einer gewissen Toleranz eine gleichbleibende Größe aufweisen. Somit ist es zulässig, auch unter Berücksichtigung der perspektivischen Abbildung, die sichtbare Textur des Ladegutes als hinreichend konstant anzunehmen. Für die Segmentierung des Ladegutes werden im Weiteren die Texturen des verbleibenden Bildbereiches analysiert.

4.3.1 Methoden der Texturklassifizierung

In vielen Bereichen besonders der industriellen Bildverarbeitung kann davon ausgegangen werden, dass sich die zu segmentierenden Bildbereiche durch homogene Intensitätsverteilung kennzeichnen und so recht einfach erkennen lassen. In diesem Fall muss jedoch davon ausgegangen werden, dass aufgrund einer variablen Beleuchtung und dem ungleichförmigen Reflektionsverhaltes des Ladegutes, die abgebildeten Objekte keine gleichmäßige Intensitätsverteilung oder Färbung zeigen. Möchte man unter diesen Voraussetzungen dennoch Bildbereiche, die einzelne Objekte zeigen, segmentieren, müssen zunächst belichtungs- und reflektionsunabhängige Merkmale, anhand derer das Objekt eindeutig identifiziert werden kann, gesucht werden.

Ein geeignetes Merkmal ist die Textur. Obwohl es für das menschliche Auge kein Problem ist, Texturen zu unterscheiden und Bildbereiche danach zu segmentieren, ist es ein nicht triviales Problem, mathematische Beschreibungen für die rechnergestützte Klassifizierung von Texturen zu formulieren. Die bestehenden Methoden für die Beschreibung von Texturen lassen sich wie folgt kategorisieren [36, 37]:

- strukturelle Betrachtung
- statistische Betrachtung
- modellbasierte Betrachtung
- transformationsbasierte Betrachtung

Die meisten strukturellen Ansätze zur Beschreibung von Texturen wurden von Haralick [38] eingeführt. Demnach können Texturen durch einen Satz von Texturprimitiven, den *microtextures*, und deren örtlicher Anordnung, den *macrotextures*, beschrieben werden. Eine Textur kann durch die Definition von Microtexturen sowie einer Funktion, welche die Wahrscheinlichkeit der örtlichen Existenz dieser Microtextur beschreibt, klassifiziert werden. Wegen der Komplexität der in der Natur auftretenden Texturen eignet sich diese Methode weniger für die Beschreibung als für die Synthese von Texturen.

Verglichen mit den strukturellen Methoden, berücksichtigen die statistischen Methoden zur Texturbeschreibung weder die Topologie noch die Geometrie [13] einer Textur (vgl. 5.2). Vielmehr werden die Verteilung sowie örtliche Beziehungen der einzelnen Intensitäts- oder Farbwerte für die Beschreibung einer Textur genutzt. Vor allem die statistischen Funktionen zweiter Ordnung bieten die Möglichkeit zur Unterscheidung Texturen [39]. Der Vorteil der statistischen Methoden erster Ordnung liegt darin, dass diese weniger aufwändig berechnet werden können. Ihr Informationsgehalt bezüglich einer Textur ist jedoch begrenzt, da sie keine Informationen über die räumlichen Beziehungen der Pixel enthalten.

Die Methoden der modellbasierten Texturbeschreibung beruhen auf der Konstruktion eines Bildmodells. Hier kommen stochastische oder fraktale Modelle zum Einsatz. Auf diese Art der Texturbeschreibung wird jedoch an dieser Stelle nicht weiter eingegangen, da sie für die weitere Datenverarbeitung nicht genutzt werden.

4.3.2 Verwendete Features und Frequenzanalyse

Mit dem Ziel, die Charakteristika der lokalen Frequenzspektren auswerten zu können, wird zu jedem Punkt im interessierenden Bildbereich eine *Fast Fourier Transformation* (FFT) des umgebenden Bildbereiches ausgeführt. Um nun ein Pixel als Ladegut identifizieren zu können, werden die Fourierkoeffizienten des transformierten Frequenzintervalls betrachtet. Die Bildbereiche, die Teile der Ladebordwand oder des Laderaumbodens zeigen, haben im Allgemeinen eine gleichmäßige Färbung und eine ebene Oberfläche. Die Textur ebener, gleichmäßig gefärbter Oberflächen kennzeichnet sich dadurch, dass ihr Leistungsspektrum den Schwerpunkt im niederfrequenten Bereich hat. Die Textur des Ladegutes hingegen ist durch eine gleichmäßige Marmorierung gekennzeichnet. Diese gleichmäßige Marmorierung kann als Signal mit dem Schwerpunkt des Leistungsspektrums in einem speziellen Frequenzbereich betrachtet werden. Somit kann durch Betrachtung der Fourierkoeffizienten zwischen den Bildbereichen, die das Ladegut zeigen und denen, die Teile des Laderaumes zeigen, unterschieden werden. Bezeichnet man den Datensatz, der ein Bild repräsentiert, als I und benennt

den Datensatz, der die Bilddaten in einer $n \times m$ Pixel großen Umgebung um das Pixel $P_{u,v}$ mit $k_{u,v}$, so ergibt sich das Leistungsspektrum $Pw_{k_{u,v}}$ der Texturen in diesem Bereich nach Gleichung 4.27.

$$\mathbf{F} = FFT\{k_{uv}\} \quad (4.26)$$

$$f_{pq} \in \mathbb{C}$$

$$Pw_{k_{u,v}} = abs(f_{pq}) \quad (4.27)$$

Hierbei bezeichnen p und q die Frequenzen, die horizontal und vertikal auftreten. Unterschreitet der mittlere Fourierkoeffizient \bar{f} im Frequenzbereich \mathbb{F} einen gegebenen Wert T_1 , so wird angenommen, dass dieses Pixel einen Teil des Ladebodens oder der Ladebordwände zeigt. Ist der mittlere Fourierkoeffizient größer als ein gegebener Wert T_1 , so wird angenommen, dass dieses Pixel zu dem Bildbereich gehört, welches das Ladegut zeigt.

$$\bar{f}_r = \frac{1}{n_{\mathbb{F}}} \sum \mathbb{F} \quad (4.28)$$

$$\mathbb{F} : \{f_{pq} \in r_{F,min} < r_F < r_{F,max}\} \quad (4.29)$$

$$r_F = \sqrt{(p^2 + q^2)} \quad (4.30)$$

$$P(u, v) = \begin{cases} 1 & \bar{f}(u, v) > T_1 \\ 0 & \bar{f}(u, v) < T_1 \end{cases} \quad (4.31)$$

Das Ergebnis der Segmentierungsfunktion ist ein Binärbild, in dem die **true** gesetzten Pixel die Bereiche markieren, die das Ladegut zeigen. Durch eine Multiplikation des Ausgangslayers mit diesem Segmentierungslayer können alle nicht interessierenden Bildbereiche gleich Null gesetzt und somit unterdrückt werden. Dieses Verfahren zur Segmentierung auf der Basis der lokalen Leistungsspektren funktioniert allerdings nur solange die aufgenommene Textur nicht zu weit vom Kamerasystem entfernt ist. Bei einer konstanten Auflösung des optischen Sensors und einer konstanten Kamerakonstante des benutzten Objektivs verschwinden mit zunehmender Entfernung zum Kamerasystem nach und nach die hochfrequenten Texturen. Dies ist darauf zurückzuführen, dass Objektpunkte mit gleicher Distanz zueinander und mit steigender Entfernung zum Kamerasystem zunehmend dicht abgebildet werden. Ab einer kritischen Entfernung der betrachteten Punkte zum Objektiv werden Texturen nur noch als annähernd homogen gefärbte Fläche abgebildet, was wiederum zur Folge hat, dass die Fouriertransformation lediglich Werte für die niedrigsten Frequenzen liefert.

5 Stereovision

Die stereoskopische Auswertung der Bilddaten stellt den eigentlichen Messvorgang des Messsystems dar. An dieser Stelle wird die Position der interessierenden Punkte im dreidimensionalen Raum \mathbb{R}^3 festgestellt. Hierfür werden sowohl die Bildbereiche, die das Ladegut zeigen, als auch die Bildbereiche, welche die Ränder zeigen, stereoskopisch ausgewertet. Die grundlegende Information für die stereoskopische Auswertung der Bilddaten ist die Disparität D der korrespondierenden Punkte. Voraussetzung hierfür ist, dass die Koordinaten der korrespondierenden Punkte bekannt sind. Die Suche nach einem korrespondierenden Punkte P_R^* im rechten Bild zu einem gewählten Punkt P_L im linken Bild wird als Korrelations- oder Pointmatching-Problem bezeichnet. Für die Lösung des Korrelationsproblems existieren unterschiedliche Ansätze. Mit Hinsicht auf die vorliegende Bildsituation müssen für die Vermessung der Ränder und die Vermessung des Ladegutes zwei unterschiedliche Verfahren für die Lösung des Korrelationsproblems eingesetzt werden. Für die Wahl eines geeigneten Verfahrens zur Lösung des Korrelationsproblems ist es notwendig, die Eigenschaften der auszuwertenden Bildbereiche zu analysieren. Die Abschnitte 5.4.1 und 5.4.3 erläutern die hier eingesetzten Verfahren zur Lösung des Korrelationsproblems. Hierfür wird in Abschnitt 5.2 eine systematische Betrachtung des für das Pointmatching relevanten Informationsgehaltes ausgeführt. Sind die Positionen der korrespondierenden Punkte bekannt, kann durch eine Triangulation der Abbildungsstrahlen die Lage des zugehörigen Objektpunktes rekonstruiert werden. Der Abschnitt 5.5 befasst sich mit der Rekonstruktion eines Objektpunktes zu einem Paar korrespondierender Punkte.

5.1 Der Prozess der Stereovision

Die Stereovisionsanalyse für diesen Prozess gliedert sich in die Verarbeitung der Bildbereiche, die die Ränder des Laderaumes zeigen und in die Bereiche, die das Ladegut zeigen. Aufgrund der unterschiedlichen Eigenschaften der Bildbereiche ist es notwendig, unterschiedliche Methoden zur Lösung des Korrelationsproblems zu nutzen. Beiden Funktionsblöcken gemeinsam ist die abschließende Triangulation der Objektpunkte im dreidimensionalen Raum. In diesem

Abschnitt wird zunächst auf die Funktion der Stereovisionsanalyse für das Ladegut eingegangen im Anschluß wird auf die Lösung des Korrelationsproblems für die Ränder eingegangen wird.

Die Stereovisionsanalyse für das Ladegut nutzt die Korrelationsfunktion, um die korrespondierenden Punkte zu den gewählten Referenzbildpunkten zu finden.

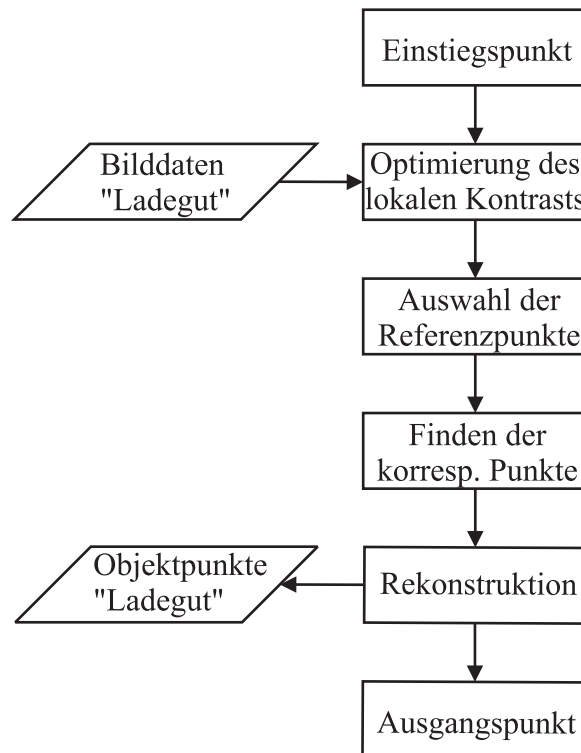


Bild 5.1: Flussdiagramm der Stereovisionsanalyse des Ladegutes

Für ein optimales Ergebnis werden zunächst lokale Belichtungsunterschiede ausgeglichen. Darüber hinaus wird der Bildbereich, der das Ladegut zeigt, hinsichtlich des lokalen Kontrastes transformiert. Das Kapitel 5.3 geht auf die Methode zur Optimierung des lokalen Kontrastes ein. Anschließend werden aus dem Datensatz des linken Bildes die Referenzpunkte ausgewählt, die hinsichtlich ihrer Eignung für das Korrelationsproblem einen maximalen Informationsgehalt aufweisen. Hierbei wird über eine Schrittweitensteuerung ein Mindestabstand der Punkte zueinander sichergestellt. Die Funktion zur Auswahl geeigneter Punkte wird in Kapitel 5.2 beschrieben. Die Liste der als geeignet eingeschätzten Bildpunkte wird dann der Funktion zur Lösung des Korrelationsproblems zugeführt. Für den Datensatz, der das Ladegut zeigt, eignet sich aufgrund der Textur des Ladegutes besonders die Auswertung

der Korrelationskoeffizienten, die sich aus der Überlagerung von Bildbereichen des linken und rechten Bildes ergeben. Das Ergebnis der Funktion für das Finden der korrespondierenden Punktpaare ist die Liste der Punktpaare sowie des dazugehörigen Korrelationskoeffizienten. Die Einzelheiten der Methode der Korrelation für das Finden der korrespondierenden Punktpaare wird in Abschnitt 5.4 erläutert. Die Liste der korrespondierenden Punktpaare wird der Funktion zur Rekonstruktion der Objektpunkte im dreidimensionalen Raum zugeführt. Hierbei werden ausschließlich die Punktpaare verwendet, deren Korrelationskoeffizient einen Mindestwert überschreiten. Das Ergebnis der Rekonstruktion ist die Liste der Punkte im Raum, die Element der Ladegutoberfläche bezüglich des Kamerakoordinatensystems sind. Die Berechnung der rekonstruierten Punkte wird in Abschnitt 5.5 beschrieben. Die Abbildung 5.1 zeigt das Flussdiagramm sowie die Ein- und Ausgabedaten der Stereovisionsanalyse für das Ladegut.

Für die Stereovisionsanalyse der Bildbereiche, die die Ränder zeigen, ist das Korrelationsproblem nicht durch Anwendung der Korrelation zu lösen.

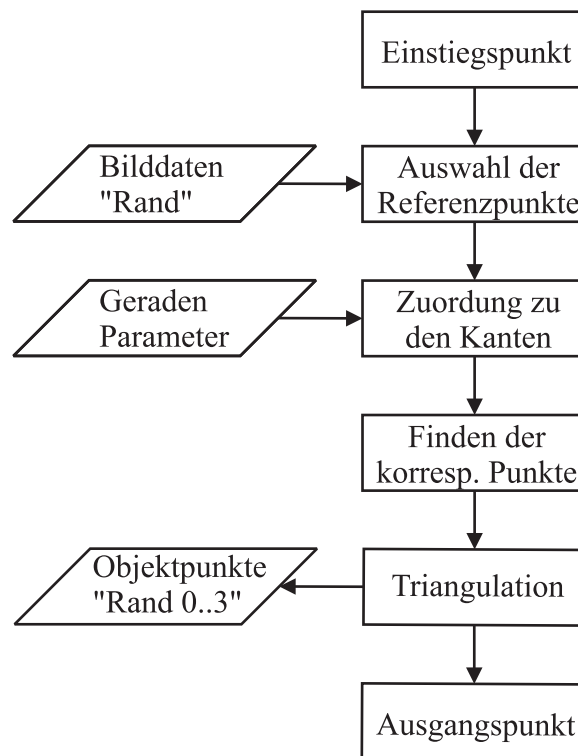


Bild 5.2: Flussdiagramm der Stereovisionsanalyse des Randes

Das Problem wird eingehend in Abschnitt 5.4.3 erläutert. Aus diesem Grund werden für

die Erfassung der korrespondierenden Randpunktpaare eine neue Funktion verwendet. Das Ergebnis dieser Funktionen ist die Liste korrespondierender Punkte, die den Rand repräsentieren. Abweichend von den Ergebnissen der Stereovisionsanalyse für die Bildbereiche, die das Ladegut zeigen, erzeugt diese Funktion keinen Wert, der auf die Qualität der Lösung schließen lässt. Die Liste der korrespondierenden Punktpaare wird dann der gleichen Funktion zur Rekonstruktion der Objektpunkte, die schon zuvor beschrieben wurde, zugeführt. Das Ergebnis ist hier die Liste der Punkte, die Element des Laderaumrandes sind. An dieser Stelle sei angemerkt, dass die Paare korrespondierender Punkte zusätzlich die Information über die Seite des Laderaumes zu der sie gehören enthalten.

5.2 Informationsgehalt eines Bildbereiches

Im Rahmen der Stereovisionsanalyse werden die Paare korrespondierender Punkte durch den Vergleich der sie umgebenden Bildbereiche ermittelt. Als ein Paar oder eine Menge korrespondierender Punkte bezeichnet man diejenigen Punkte, die von einem Objektpunkt aus dem dreidimensionalen Raum \mathbb{R}^3 in mehreren Kameras abgebildet werden. Aufgrund der Sensoranordnung darf hier angenommen werden, dass die Basislänge $|\vec{V}_B| = 140 \text{ mm}$ des Epipolar-systems im Vergleich zur Entfernung des abgebildeten Objekt $1000 \text{ mm} < Z < 3000 \text{ mm}$ sehr klein ist, sodass die Bildbereiche um die Paare korrespondierender Punkte in den beteiligten Kameras hinreichend gleich abgebildet werden, wodurch eine eindeutige Zuordnung zweier Punkte zu einem Paar korrespondierender Punkte aufgrund ihrer umgebenden Bildbereiche möglich wird. Die Lösung des Korrelationsproblems durch den Vergleich von Bildbereichen setzt voraus, dass diese Bildbereiche in der Umgebung der betrachteten Bildpunkte hinreichende Informationen für eine eindeutige Lokalisierung dieses Punktes in der Ebene enthält. Somit kann das relevante Merkmal für den Informationsgehalt als die Eigenschaft einer Textur, sich in einem Bildbereich eindeutig zuordnen zu lassen, definiert werden. Hierfür ist sowohl die Intensität der abgebildeten Textur als auch die Art des abgebildeten Musters ausschlaggebend.

Im Allgemeinen können die Topologie und die Geometrie einer Bildpunktemenge nach W. Schroeder et.al. [13] für die lokale Beziehung von Bildpunkten zueinander betrachtet werden. Die Topologie einer Punktemenge beschreibt eine Klasse geometrischer Objekte, der die Punkte in dieser Kombination zugeordnet werden können, wobei die Topologie einer Punktemenge noch keine Information über die konkreten Entfernungen der Punkte zueinander enthält. Sie sagt lediglich etwas über die grundsätzliche Anordnung der Punkte zueinander aus. Die Geometrie einer Punktgruppe hingegen kann als Instanzierung einer Topologie

angesehen werden. Abbildung 5.3 verdeutlicht den Zusammenhang zwischen der Topologie einer Punktemenge und einer konkreten Geometrie dieser Topologie.

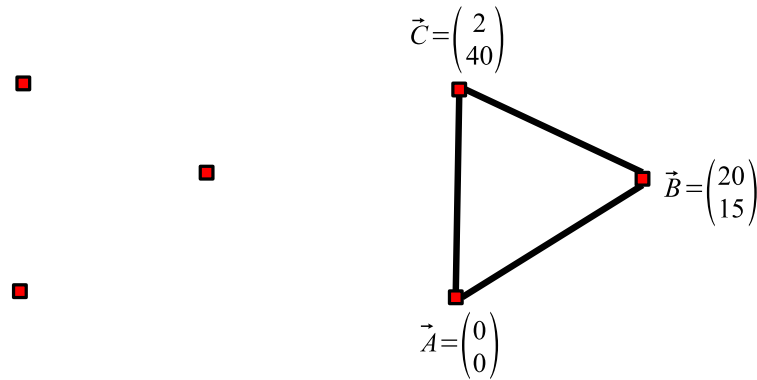


Bild 5.3: Systematische Betrachtung von Punktekongstellationen

Während die links dargestellten Punkte ein beliebiges Dreieck ohne Angabe von Drehsinn oder Längenverhältnissen repräsentieren, wird durch die Einordnung der Punkte in ein einheitliches Koordinatensystem sowie die Bezeichnung der Punkte ein Dreieck explizit definiert.

Neben der räumlichen Anordnung der Pixel des betrachteten Merkmals ist die abgebildete Textur sowie deren Intensität an dieser Stelle von zentraler Bedeutung. Betrachtet man nun nicht mehr die Pixel eines Merkmals sondern alle Pixel einer Menge von Merkmalen sowie deren Geometrien, so kann dies als eine Textur angesehen werden.

Die Verteilung der Intensität der abgebildeten Textur entspricht dann der Nutzung des darstellbaren Wertebereichs und wird im Allgemeinen als Kontrast dieses Bildbereiches angesehen. Der Kontrast der abgebildeten Textur beeinflusst seinerseits den Wertebereich des Ergebnisses der Korrelationsfunktion zweier überlagerter Bildbereiche und somit den Kontrast des Ergebnisses. Die Identifikation eines Extremums wird um so einfacher und präziser, je deutlicher es sich vom Mittelwert der Werte des betrachteten Wertebereiches abhebt. Für die Ermittlung des korrespondierenden Punktes ist also ein hohes Maß an Kontrast der Korrelationsfunktion und somit der betrachteten Wertebereiche von Interesse. Der Kontrast einer Wertemenge kann durch die Standardabweichung bzw. die Varianz oder die Entropie mathematisch beschrieben werden.

Darüberhinaus beeinflusst die Art der Textur, also das sichtbare Muster, den Informationsgehalt eines Bildbereichs. Die Abbildung 5.4 läßt erkennen, dass der markierte Bildbereich entlang der abgebildeten Linie verschoben werden kann, ohne dass sich das Ergebnis der Korrelationsfunktion ändert und somit keine eindeutige Zuordnung zu einem Bildbereich zulässt. Der Informationsgehalt in dieser Richtung ist also nicht ausreichend.

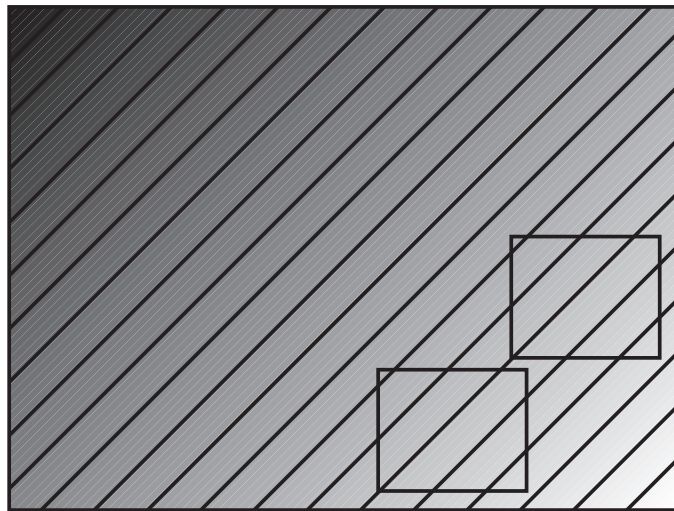


Bild 5.4: Informationsgehalt für das Korrelationsproblem

Die Eigenschaft einer Textur lokal eindeutig bestimmt zu werden, kann durch die örtliche Relation von Pixeln mit gleicher numerischer Merkmalsausprägung beschrieben werden. Diese Eigenschaft kann entsprechend Abbildung 5.3 als Topologie einer Textur beschrieben werden. Die Abbildung 5.5 zeigt die topologische Ebene der schwarz markierten Pixel. Je nach Farbauflösung und Ausnutzung des zur Verfügung stehenden Wertebereichs entstehen so mehr oder weniger viele topologische Ebenen. Die Topologie einer Textur ist die Basis für die Entscheidung darüber, ob ein ausgewählter Bildbereich genau einer Position in einem Bild zugeordnet werden kann.

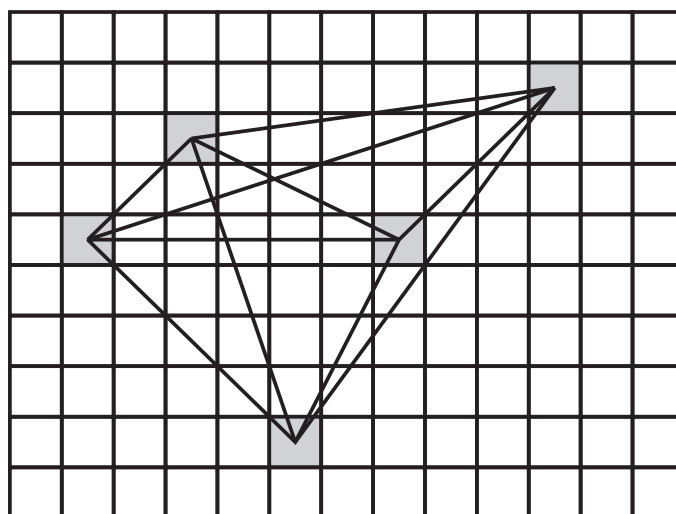


Bild 5.5: Topologie einer Textur

In Abbildung 5.4 wird deutlich, dass die Topologie des abgebildeten Streifenmusters nicht geeignet ist, den markierten Kernel eindeutig zu lokalisieren. Auch eine Variation des Kontrasts schafft an dieser Stelle keine Abhilfe. Betrachtet man hingegen Abbildung 5.6 so wird deutlich, dass hier eine eindeutige Zuordnung des markierten Kernels möglich ist.

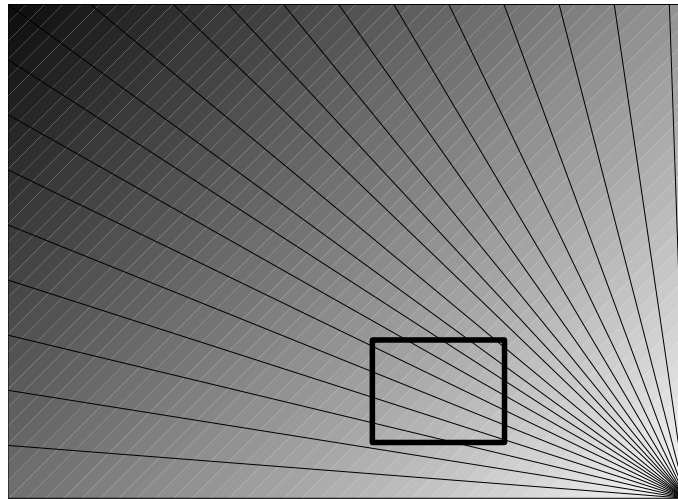


Bild 5.6: hinreichender Informationsgehalt eines Bildausschnittes

Um nun Punkte auszuwählen, die sich besonders für die Anwendung der Korrelationsfunktion eignen, müssen die sie umgebenden Bildbereiche hinsichtlich des Kontrastes und der Topologie ihrer Textur geprüft werden. Wie bereits erwähnt, eignen sich statistische Methoden wie die Variation oder die Entropie, um den Kontrast des betrachteten Bildbereiches zu bewerten. Allerdings geben diese Methoden keinen Aufschluss über die Topologie der Textur. Die markierten Bildbereiche aus den Abbildungen 5.4 und 5.6 ergeben eine annähernd gleiche Varianz und Entropie über den durch die Markierungen festgelegten Wertebereich, obwohl der Bildausschnitt aus Abbildung 5.6 örtlich definiert ist und der Bildausschnitt aus Abbildung 5.4 nicht. Um die Topologie einer Textur bewerten zu können, ist es notwendig, den Informationsgehalt entlang bestimmter Richtungen im Bild zu bestimmen. Die Wertevarianz der Pixelreihen parallel zu den in Abbildung 5.4 sichtbaren Linien ergibt sich offensichtlich zu *NULL*. Dieses ist die Richtung, in der dieser Bildbereich nicht eindeutig lokalisiert werden kann. Eine geringe Varianz in einer Richtung deutet auf einen Mangel an Information für die Lokalisierung des Bildbereiches entlang dieser Richtung hin. Um nun einen Punkt und dessen umgebenden Bildbereich hinsichtlich seiner Eigenschaft der lokalen Eindeutigkeit bewerten zu können, liegt es nahe, die Varianz in allen Richtungen strahlförmig um den Bezugspunkt zu bilden und den kleinsten Wert aus dieser Menge als relevanten Wert zu verwenden. Mit

Hinsicht auf den Berechnungsaufwand werden die Varianzen der Pixelwerte entlang der Symmetrielinien und der Hauptdiagonalen ausgewertet. Da es sich bei der betrachteten Textur um ein stochastisches Muster handelt, ist dies hinreichend. Darüber hinaus ist die Berechnung der Varianz bei gleichem Informationsgehalt weniger aufwändig als die der Entropie. Aus diesem Grund wurde an dieser Stelle die Varianz zur Bewertung des lokalen Informationsgehaltes gewählt. In Abbildung 5.7 sind die Pixel des Bildbereiches markiert, die zur Bestimmung des lokalen Informationsgehaltes herangezogen wurden.

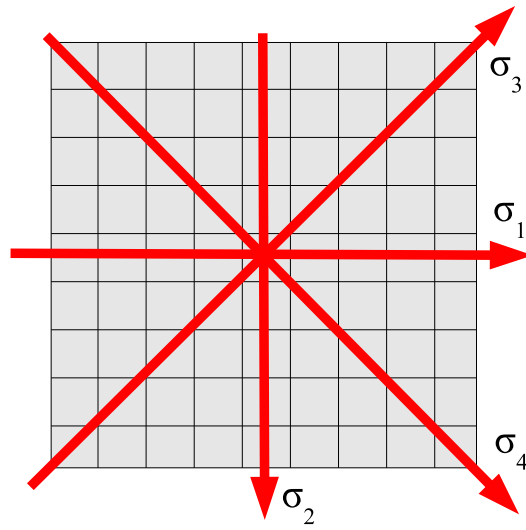


Bild 5.7: Berechnung des Informationsgehaltes eines Bildbereiches

$$IC_{uv} = \min_{j=0}^n (\sigma_j)_{uv} \quad (5.1)$$

$$\sigma_j = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

Hierbei bezeichnet IC_{uv} den lokalen Informationsgehalt eines Bildausschnittes an der Stelle u, v . σ_j benennt die Varianz und \bar{x} den Mittelwert der Pixelwerte entlang der in Abbildung 5.7 markierten j -ten Richtung. Abschließend sei angemerkt, dass dieses Verfahren auch auf periodische Texturen angewendet werden kann. Hierbei ist es notwendig, die Größe der betrachteten Bildausschnitte und der zu untersuchenden Umgebung so zu wählen, dass Phasenmehrdeutigkeiten nicht berücksichtigt werden müssen.

5.3 Bildvorverarbeitung

Im vorhergehenden Abschnitt 5.2 wurde bereits auf die Bedeutung des Kontrastes einer Textur hinsichtlich ihrer Eignung für das Pointmatching beschrieben. In diesem Abschnitt wird nun eine Methode erläutert, die dem Ausgleich lokaler Kontrastdefizite dient.

Aufgrund der natürlichen Beleuchtungsverhältnisse unterliegen die eingezogenen Bilder Schatteneinflüssen. Davon sind besonders die Bereiche des Ladegutes betroffen, die dicht an der Ladebordewand gelegen sind. Durch die in Schattenbereichen reduzierte Beleuchtung der betrachteten Szene wird auch der Kontrast dieser Bereiche maßgeblich reduziert. Um auch an diesen Stellen noch ein hinreichend aussagekräftiges Ergebnis der Korelationsfunktion zu erhalten, wird versucht, auf der Bildebene den Schatteneinfluss zu reduzieren und darüberhinaus den lokalen Kontrast zu verbessern. Für die Unterdrückung von lokalen Beleuchtungsvarianzen wird an dieser Stelle der *HighBoost-Filter* [54] eingesetzt. In Anbetracht der Tatsache, dass der Schatten im Vergleich zur sichtbaren Textur eine niederfrequente Störung des Signals darstellt, kann der HighBoost-Filter zur Klasse der Hochpassfilter gezählt werden. In den Bereichen, die von einem Schatten überlagert sind, liegt offensichtlich eine niedrigere durchschnittliche Helligkeit vor als in den Bereichen, die voll ausgeleuchtet sind. Solange die abgeschatteten Bereiche nicht vollständig abgedunkelt sind, bleibt das Textursignal, wenn auch mit geringerem Kontrast erhalten. Die Abbildung 5.8 verdeutlicht den Einfluss des Schattens auf die Pixelwerte anhand einer Pixelzeile.

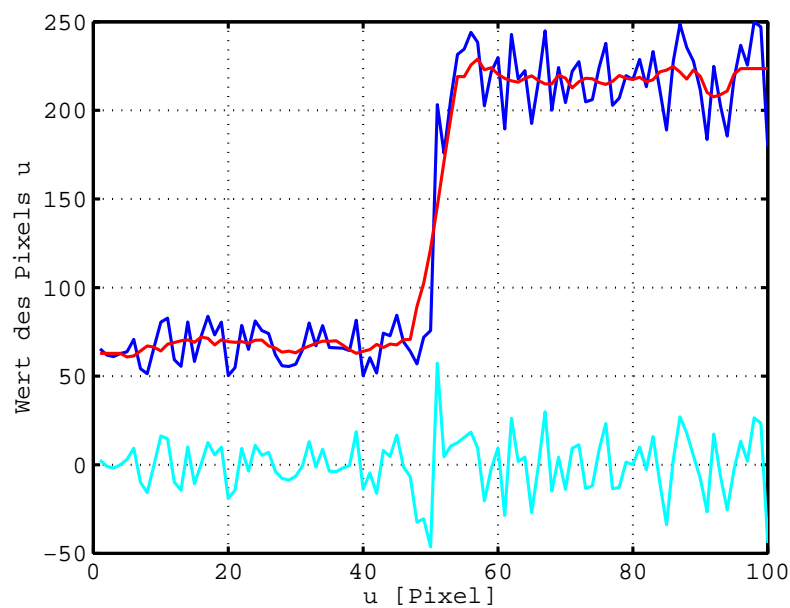


Bild 5.8: Signal einer Pixelzeile mit einem abgeschatteten Bereich

Die blaue Linie kennzeichnet die Pixelwerte einer Zeile des Bildes. Die linke Hälfte des Datensatzes gehört zu einem abgeschatteten Bildbereich während die rechte Hälfte zu einem voll ausgeleuchteten Bildbereich gehört. Man sieht, dass der mittlere Pixelwert in der linken Hälfte deutlich kleiner ist als der mittlere Pixelwert auf der rechten Hälfte des Datensatzes. Bildet man nun den gleitenden Mittelwert über die Pixelwerte, so erhält man die in Abbildung 5.8 rot eingezeichnete Linie. Die cyangefärbte Linie ergibt sich durch Subtraktion der roten Linie von der blauen Linie. Man erkennt, dass die hochfrequente Information weitestgehend erhalten bleibt. Der Datensatz wurde lediglich mittelwertfrei gemacht. Die interessierende Information der Textur bleibt also weitestgehend unverändert erhalten, während der Schatteneinfluss reduziert wurde. Es verbleibt der Effekt, dass durch den Schatten der Kontrast der Textur reduziert wird. Dies wird daran sichtbar, dass die in Abbildung 5.8 cyangefärbte Linie in der linken Hälfte immer noch eine geringere mittlere Amplitude als in der rechten Hälfte aufweist. Es kommt hinzu, dass der Wertebereich der mittelwertfreien Datenreihe auch negative Werte einschließt. Negative Pixelwerte werden aber üblicherweise nicht ohne Weiterverarbeitung in der digitalen Bilddarstellung verwendet.

Um den Kontrast im abgeschatteten Bereich sowie im ausgeleuchteten Bereich auf ein gleiches optimales Maß zu erhöhen, wird abschnittsweise das Signal auf den maximalen Wertebereich normiert. Die Abbildung 5.9 zeigt das Ergebnis der Optimierung der Bildqualität am Beispiel einer Pixelzeile.

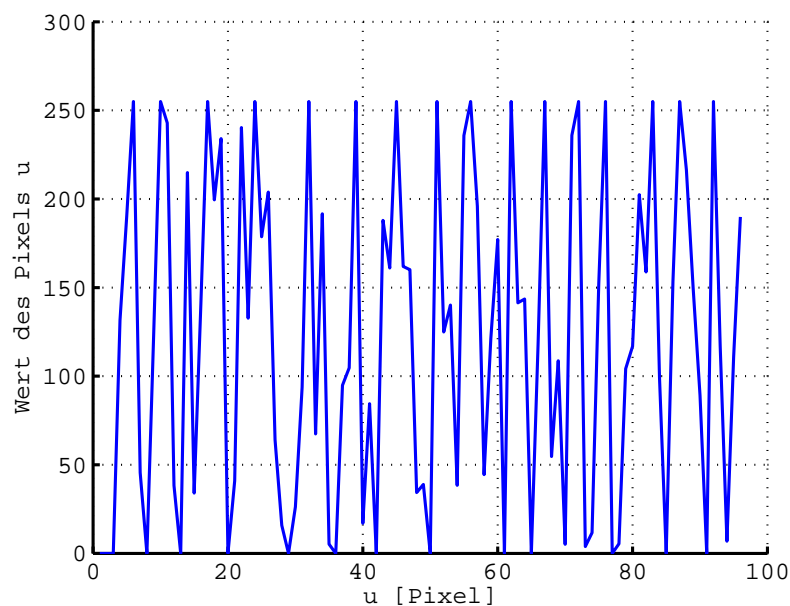


Bild 5.9: Normiertes Signal einer Pixelzeile

Im Weiteren soll die Implementierung des hier angeführten Verfahrens für die zweidimensionalen Datensätze der Bilder erläutert werden. Mit Hinsicht auf den Berechnungsaufwand ist es sinnvoll, die Filteralgorithmen in Form linearer Standardfilter abzubilden, da diese mit Hilfe des Faltungssatzes besonders schnell zu berechnen sind. Den gleitenden Mittelwert über die Pixelwerte eines Bildes kann man durch die Faltung des Bildes mit einem Gauß-Tiefpassfilter annähern. Bezeichnet man einen zweidimensionalen Gauß-Tiefpassfilter mit h_{GTP} , so ergibt sich das Bild I_{TP} , über welches der angenäherte gleitende Mittelwert gebildet wurde, nach Gleichung 5.2.

$$I_{TP} = I * h_{GTP} \quad (5.2)$$

I bezeichnet dabei das Eingabebild und $*$ den Faltungsoperator. Das HighBoost gefilterte Bild I_{HB} ergibt sich dann durch Subtraktion des tiefpassgefilterten Bildes vom Ausgangs-Bild nach Gleichung 5.3.

$$I_{HB} = I - I_{TP} \quad (5.3)$$

Die Bildbereiche, die für die Triangulation markiert wurden, zeigen das Ladegut. Sieht man von der entfernungsabhängigen zentrischen Streckung ab, kann davon ausgegangen werden, dass in diesem Bildbereich eine annähernd gleich feine Textur sichtbar ist. Betrachtet man den Kontrast eines Bildbereiches als Amplitude des Signals der Pixelwerte, so kann man den lokalen Kontrast als Leistung des Signals interpretieren. Vor dem Hintergrund, dass eine gleichbleibende Textur auch ein gleichbleibendes Frequenzspektrum mit sich bringt, ist die so definierte Leistung nur noch von der Amplitude dieses Signals abhängig. Die Leistung dieses Signals kann also als Maß für den Kontrast eines Bildbereiches genutzt werden. Somit kann durch Mittelung der absoluten Pixelwerte eines Bildbereiches dessen Kontrast bemessen werden. Die Faltung ist eine gewichtete Summation der Elemente des ausgewählten Bereiches. Durch eine geeigneten Gewichtung der Elemente kann der Mittelwert des betrachteten Bildausschnittes berechnet werden. Somit ist es möglich, mit Hilfe der Faltung eines Bildes den gleitenden Mittelwert über die Pixelwerte anzunähern. Nach Graefe/Feitosa [1] kann demnach ein Maß für den lokalen Kontrast durch Gleichung 5.4 berechnet werden.

$$I_{lc} = h_{GTP} * |I_{HB}| \quad (5.4)$$

Hierbei bezeichnet I_{lc} den Datensatz, der für jedes Pixel ein Maß des Kontrasts seiner Umgebung angibt. Abschließend wird das Maß des lokalen Kontrastes auf den Wertebereich von 0 bis 1 normiert.

$$I_n = \begin{cases} \frac{h_{hp}}{I_{lc}} & I_{lc} > 0 \\ 0 & I_{lc} = 0 \end{cases} \quad (5.5)$$

Die Abbildung 5.10 zeigt links ein segmentiertes Bild und rechts das Bild nach der Optimierung des lokalen Kontrastes.

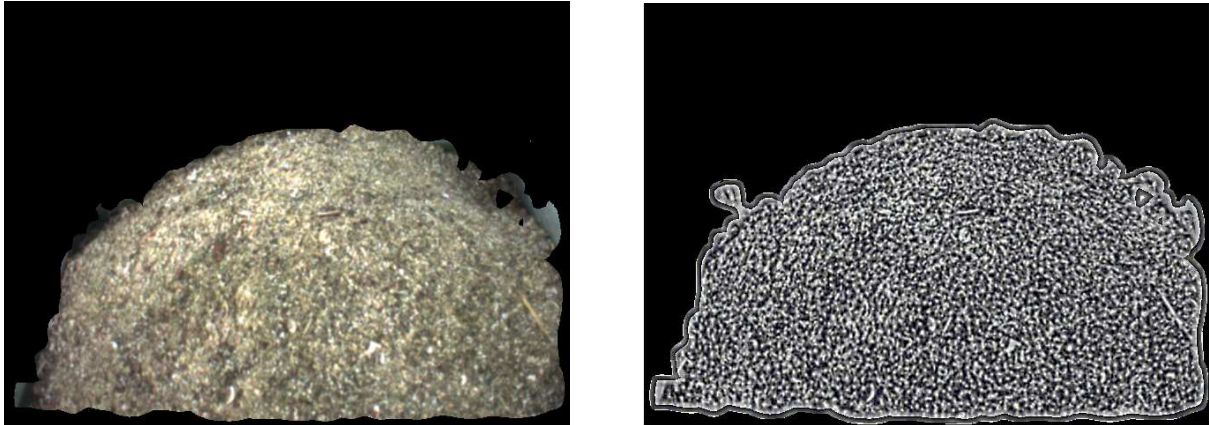


Bild 5.10: Bild nach Optimierung der Bildqualität

Es fällt auf, dass die jetzt sichtbare Textur in allen Bildbereichen des Ladegutes einen annähernd gleichen Kontrast aufweist, während die Textur des Ladegutes im linken Bild an den Rändern des segmentierten Bereiches im Schatten der Ladebordwand liegt. Durch die Transformationen ist die Textur als solches nicht mehr sichtbar. Für die weitere Verarbeitung der Bilder im Rahmen des Pointmatchings ist das aber nicht weiter von Belang. Für das in diesem Fall eingesetzte Verfahren zur Lösung des Korrelationsproblems ist es lediglich nötig, dass ein Bildbereich aus dem linken Bild mit genau einem Bildbereich aus dem rechten Bild korreliert. Die Details hierzu werden im folgenden Abschnitt 5.4.1 erläutert.

5.4 Die Lösung des Korrelationsproblems (Pointmatching)

Um den Füllstand zu bemessen, soll die Oberfläche des Ladegutes im \mathbb{R}^3 modelliert werden. Hierfür wird mit der Kenntnis der Lage korrespondierender Punktepaaare, sowie der bekannten Geometrie des Kamerasystems die Entfernung des zugehörigen Objektpunktes ermittelt. Es stellt sich das Kernproblem zu einem ausgewählten Punkt P_L des linken Bildes den korrespondierenden Punkt P_R^* im rechten Bild, der das selbe Objekt abbildet, zu finden. Für die Auswertung ist es nötig, einerseits die Lage der Ränder des Laderaumes als auch die Lage der Schüttgutoberfläche im Weltkoordinatensystem zu erfassen. Da die Bildbereiche, die das Ladegut zeigen, grundsätzlich anders charakterisiert sind als diejenigen, die den Rand zeigen, muss für die Suche nach den korrespondierenden Punkten für beide Merkmale ein geeignetes Verfahren gefunden werden. Im folgenden Abschnitt 5.4.1 wird auf die Bildbereiche,

die das Ladegut zeigen, die Korrelationsfunktion für die Lösung des Korrelationsproblems angewendet. In Abschnitt 5.4.3 wird ein merkmalsbasiertes Verfahren zur Bestimmung des korrespondierenden Punktes auf die Bildpunkte, die dem Objekt *Rand* zugeordnet werden, angewandt.

5.4.1 Das Pointmatching für das Ladegut

Wie schon erwähnt, weist das Ladegut eine stochastische Textur auf. Somit kann davon ausgegangen werden, dass sich in einem begrenzten Bildbereich die Textur nicht wiederholt. Diese Tatsache ermöglicht es, einen Punkt anhand des ihn umgebenden Bildbereiches eindeutig zu lokalisieren. Darüberhinaus ist die Basislänge des Epipolarsystems sehr viel kleiner als die Entfernung der Kameras zum Messobjekt. Zusätzlich sind die Kameras annähernd achsparallel angeordnet. Es kann also davon ausgegangen werden, dass keine nennenswerten perspektivischen Verzerrungen auftreten, sodass die Kameras die Bilder, abgesehen von einem deutlichen translativen Versatz der Szene, annähernd identisch einziehen. Somit ist es möglich, die korrespondierenden Punkte durch einen Vergleich der umgebenden Bildbereiche zu identifizieren. Für den Vergleich der Bildbereiche wird die Korrelationsfunktion genutzt. Der Korrelationskoeffizient zweier Bildbereiche gibt Aufschluss über das Maß der Übereinstimmung der Bildbereiche. Somit wird als korrespondierender Punkt P_R^* zu einem Referenzpunkt P_L im linken Bild der Punkt gewählt, dessen Umgebung einen maximalen Korrelationskoeffizienten mit dem Referenzbildausschnitt aufweist. Definiert man die Umgebung N_P um ein Pixel P mit den Dimensionen $f \times g$, so kann man die normierten Korrelationskoeffizienten dieses Bildbereichs mit einem anderen Bildbereich der gleichen Dimension durch Gleichung 5.6 berechnen.

$$c = \frac{\sum_f \sum_g (a_{fg} - \bar{a})(b_{fg} - \bar{b})}{\sqrt{\sum_f \sum_g (a_{fg} - \bar{a})^2 \sum_f \sum_g (b_{fg} - \bar{b})^2}} \quad (5.6)$$

Hierbei bezeichnet a_{fg} bzw. b_{fg} den Wert eines Pixels an der Stelle f, g im gewählten Bildbereich. \bar{a} und \bar{b} bezeichnen den Mittelwert über alle Pixelwerte des gewählten Bildbereichs nach Gleichung 5.7.

$$\bar{a} = \frac{1}{fg} \sum_f \sum_g a_{fg} \quad (5.7)$$

Definiert man um einen Punkt P_{Ruv} im rechten Bild einen Bildbereich N_R mit der Größe $k \times l$ mit $k > f$ und $l < g$ so kann zu jedem Punkt P_{fg} aus N_R des rechten Bildes der normierte

Korrelationskoeffizient $c(P_{fg})$ $P_{fg} \in N_R$ des ihn umgebenden Bildbereiches $N_{P_{fg}}$ mit dem Referenzbildbereich $N_{P_{Lm}}$ berechnet werden.

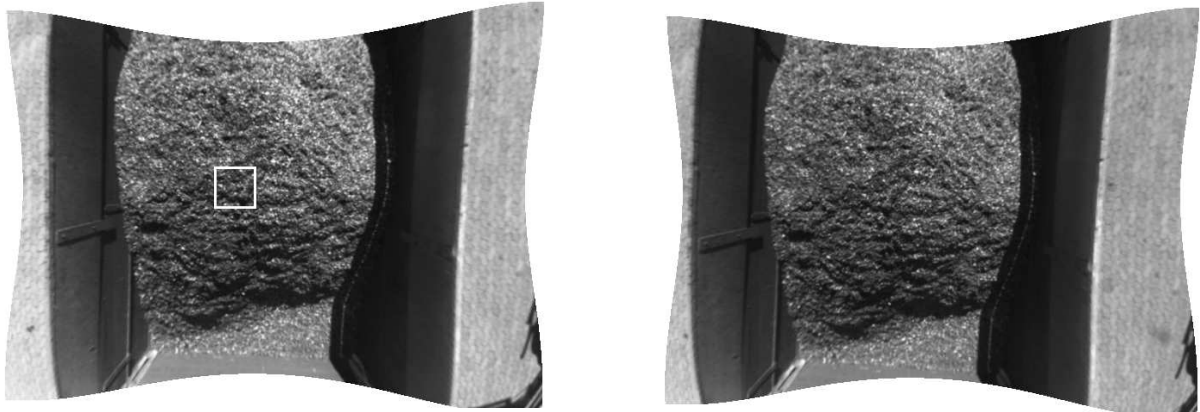


Bild 5.11: Rektifiziertes Stereobildpaar des Laderaumes

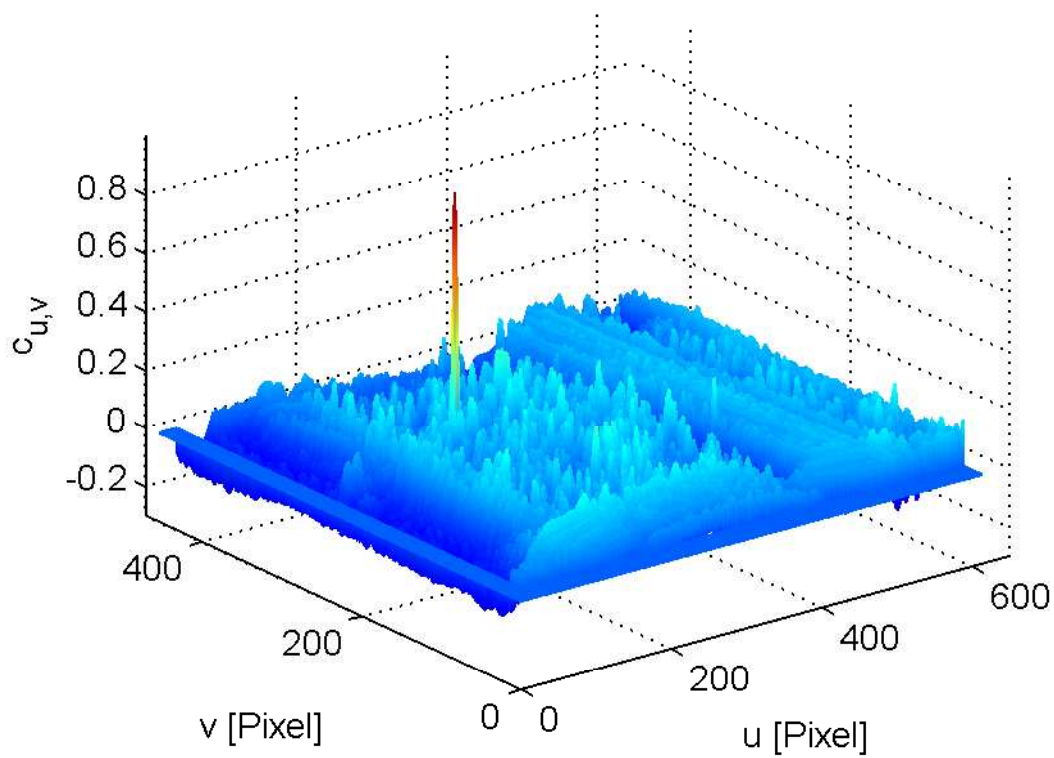


Bild 5.12: Matrix der normierten Korrelationskoeffizienten

Die Ergebnisse werden in der Matrix C_m abgelegt. Die Matrix C_m wird als Matrix der normierten Korrelationskoeffizienten bezeichnet und hat die Größe $(k - f) \times (l - g)$. Die Matrix C_m gibt also Aufschluss über die Übereinstimmung der Bildbereiche des rechten Bildes mit dem Referenzbildausschnitt. Unter der Voraussetzung, dass sich der korrespondierende Punkt P_R^* des rechten Bildes zu einem Punkt P_L aus dem linken Bild durch optimal Übereinstimmung der umgebenden Bildbereich auszeichnet, kann durch Auswahl des Punktes, der den maximalen normierten Korrelationskoeffizienten liefert, der optimal korrespondierende Punkt bestimmt werden. Die Abbildung 5.12 zeigt die Matrix der normierten Korrelationskoeffizienten für den rot markierten Referenzbildausschnitt. Man erkennt, dass für einen Punkt ein deutlich größerer normierter Korrelationskoeffizient auftritt. Hier kann der korrespondierende Punkt P_R^* angenommen werden. Der korrespondierenden Punkt P_R^* zu dem Punkt P_L ergibt sich aus der Bedingung 5.8.

$$P_R^* = \operatorname{argmax}(C_m) \quad (5.8)$$

Dieses globale Maximum deutet auf den Bildbereich hin, der die beste Übereinstimmung mit dem Referenzbildausschnitt aufweist. Um nun den korrespondierenden Punkt P_R^* zu einem Referenzpunkt P_L zu finden, wird die Matrix der normierten Korrelationskoeffizienten C_m zu einem Punkt P_l gebildet und deren Maximum als optimale Lösung des Korrelationsproblems betrachtet.

Da das zu vermessende Objekt in einem begrenzten Entfernungsbereich zum Kamerasystem liegt, wird dementsprechend die Disparität der korrespondierenden Punkte auch auf einen begrenzten Wertebereich beschränkt sein. Die Suche nach einem korrespondierenden Punkt kann also auf die nähere Umgebung um die Koordinate des Referenzpunktes reduziert werden. Darüberhinaus ist die Position der Kameras zueinander fest, sodass nur Disparitäten mit einheitlichem Vorzeichen auftreten können. Was die Menge \mathbb{P}_R^* der Kandidaten für die korrespondierenden Punkte weiter einschränkt. Die schärfste Restriktion für die Gültigkeit eines korrespondierenden Punktes zu einem Referenzpunkt ist die Zugehörigkeit zur Epipolarlinie des Referenzpunktes. Nach Gleichung 2.31 muss gemäß Emanuele Trucco [10] ein korrespondierender Punkt auf der Epipolarlinie des jeweils zugehörigen Referenzpunktes liegen. Somit kommen nur noch die Punkte als korrespondierende Punkte in Frage, die in einem Segment der Epipolarlinie liegen, welches durch die Koordinaten des Referenzpunktes bestimmt werden kann.

5.4.2 Die Bedeutung der Rektifizierung

Um die Gleichheit zweier Bildbereiche mit Hilfe der Korrelationsfunktion zu untersuchen, ist es besonders wichtig, dass die Bildbereiche auf koplanare und achsparallele Bildebenen abgebildet wurden.



Bild 5.13: Korrelationskoeffizient bei rotierten Bildern

Der Korrelationskoeffizient nach Gleichung 5.6 bewertet die Einträge der betrachteten Datenfelder bezüglich ihres Index. Da durch einen rotatorischen Versatz jedoch die Lage der korrespondierenden Punkte in den Koordinatensystemen der Bildebenen nicht gleich ist, wird der Kontrast im Lösungsdatensatz des Korrelationsproblems schwächer, was wiederum die Identifizierung des korrespondierenden Punktes erschwert. Die Abbildung 5.13 verdeutlicht das Problem. Die Abbildung 5.13 zeigt rechts und links das gleiche Bild mit dem Unterschied, dass das rechte Bild um einen Winkel von $\phi = 0,1$ um das Zentrumspixel $P_z = [320 \ 240]^T$ des Adressbereiches gedreht wurde. Es ist offensichtlich, dass der normierte Korrelationskoeffizient des blau eingerahmten Bildbereiches für auf sich selbst angewendet genau gleich $c_{P_z} = 1$ ist. Der normierte Korrelationskoeffizient für den markierten Bildbereich im linken Bild und den Bildbereich um den korrespondierenden Punkt des gedrehten Bildes ergibt sich aber nur noch zu $c_{P_z, \text{rot}} = 0,508$. Es wird deutlich, wie wichtig die koplanare und achsparallele Anordnung der Bilder für die Suche nach den korrespondierenden Punkten ist. Gleiches gilt für ungleichförmige Verzeichnung der Bilder durch die Objektive oder unterschiedlich starke zentrische Streckung der Bilder durch die Abbildung der Szene auf nicht koplanare Bildebenen. Aus diesem Grund werden die Bilder wie in Kapitel 3.5 beschrieben rektifiziert. Durch die Rektifizierung werden die Bilder so transformiert, dass sie der Abbildung durch zwei koplanar angeordnete, achsparallele, nur entlang der X -Achse verschobene, ideale Kameras entsprechen. Unter dieser Voraussetzung ergibt sich die Epipolarlinie l_P im rechten

Bild zu einem beliebigen Punkt P_L im linken Bild nach Gleichung 2.50 zu einer horizontalen Linie. Stellt man die Essentialmatrix \mathbf{E}_{rect} nach Gleichung 2.41 mit den Werten aus Tabelle 3.4, so ergibt sich \mathbf{E}_{rect} für das rektifizierte Bildpaar zu:

$$\begin{aligned}\mathbf{E}_{\text{rect}} &= [\mathbf{R}_{\text{rect}} \overrightarrow{t_{\text{rect}}}] \\ \mathbf{R}_{\text{rect}} &= \mathbf{I} \\ \mathbf{E}_{\text{rect}} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -x_0 \\ 0 & x_0 & 0 \end{pmatrix}\end{aligned}\quad (5.9)$$

Hierbei bezeichnet \mathbf{I} eine 3×3 Einheitsmatrix. Somit ergibt sich die Epipolarlinie l_P zu einem Punkt P nach Gleichung 5.10.

$$\overrightarrow{l_p} = \mathbf{E}_{\text{rect}} \overrightarrow{P} \quad (5.10)$$

$$\overrightarrow{l_p} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -x_0 \\ 0 & x_0 & 0 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix} \quad (5.11)$$

$$\overrightarrow{l_p} = \begin{pmatrix} 0 \\ -x_0 \\ P_y x_0 \end{pmatrix} \quad (5.12)$$

Nach Gleichung 2.42 lautet die Gleichung der Epipolarlinie in Normalform:

$$l_p : 0 = y - P_y \quad (5.13)$$

Für ein rektifiziertes Bildpaar ist die Epipolarlinie zu einem beliebigen Punkt eine horizontale Gerade für die gilt $y = P_y$. Dies vereinfacht die Suche nach den korrespondierenden Punkten erheblich. So wird nicht nur eine deutliche Reduzierung der zu untersuchenden Punkte im rechten Bild erreicht, sondern auch eine wesentliche Reduzierung von falschen korrespondierenden Punkten. Im idealen Fall würde sich also die Menge der in Frage kommenden Punkte auf die Punkte eines Geradenabschnittes beschränken. Aufgrund von Rundungsfehlern und durch die Abbildung auf eine diskrete Bildebene müssen aber die Abbildungsfehler berücksichtigt werden. Für die Suche nach dem korrespondierenden Punkt P_R^* zu einem Referenzpunkt P_L wird hier ein Streifen mit einer Breite von drei Pixeln und einer Länge von $\pm \Delta D$ Pixeln um die mittlere erwartete Disparität \bar{D} gewählt.

Die Abbildung 5.4.2 zeigt den Bildbereich des rechten Bildes, in dem nach dem korrespondierenden Punkt gesucht wird.

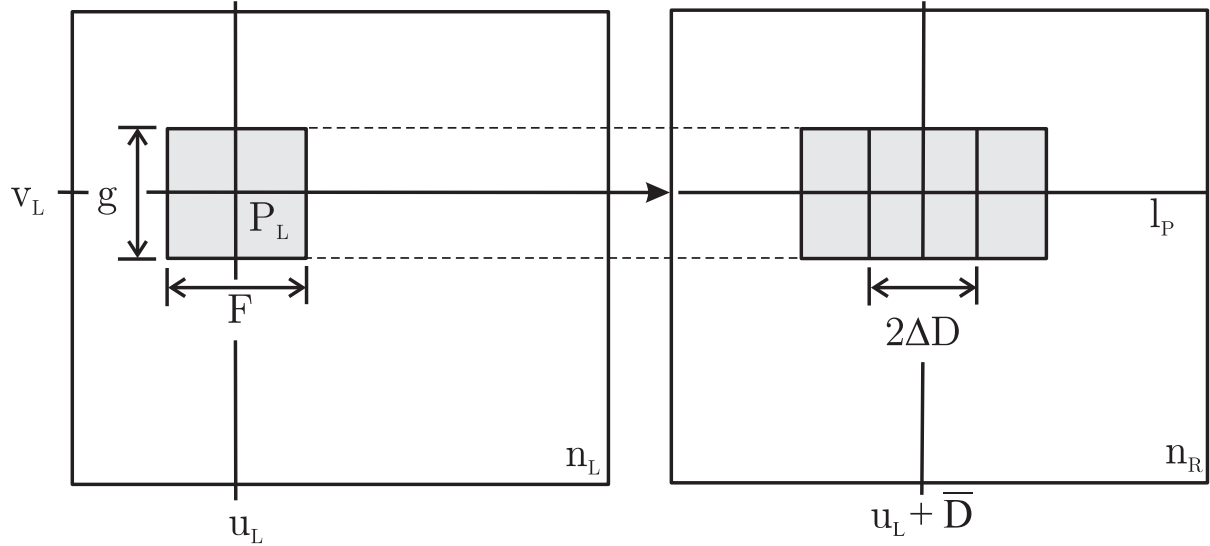


Bild 5.14: Suchfenster für den korrespondierenden Punkt

Aufgrund der Anordnung der Kameras kann im Vorfeld das erwartete Vorzeichen bestimmt, sowie die Größenordnung der Disparität abgeschätzt werden. Nach Gleichung 2.32 gilt für den Zusammenhang der korrespondierenden Punkte P_L und P_R^* zu einem Punkt P

$$P_R^* = \mathbf{H}_{LR} P_L$$

Für die hier verwendeten rektifizierten Bilder kann nach Kapitel 3.5.1 angenommen werden, dass sie mit achsparallelen und koplanar angeordneten Kameras, die keinen Abbildungsfehler durch die Linse erzeugen, aufgenommen wurden. Somit ergibt sich die Homographiematrix H_{LR} für die Abbildung des Punktes P_R^* in das Koordinatensystem der linken Kamera nach 2.34.

$$\mathbf{H}_{LR} = \mathbf{A}_R [\mathbf{R}_R \vec{t}_r] \mathbf{A}_L^{-1}$$

Für diesen Fall reduzieren sich die Abbildungs- und Transformationsmatrizen zu:

$$\mathbf{A}_L = \mathbf{A}_R = \begin{pmatrix} -c & 0 & 0 \\ 0 & -c & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Die Rotationsmatrix \mathbf{R}_R ist dann gleich der Einheitsmatrix und der Translationsvektor \vec{t}_r repräsentiert nur den Versatz der Kameras entlang der X Achse.

$$\mathbf{R}_R = \mathbf{I}$$

$$\vec{t}_r = \begin{pmatrix} -x \\ 0 \\ 0 \end{pmatrix}$$

Somit ergibt sich die Homographiematrix \mathbf{H}_{RL} zu

$$\begin{aligned} \mathbf{H}_{RL} &= \begin{pmatrix} -c & 0 & 0 \\ 0 & -c & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -x \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} -c & 0 & 0 \\ 0 & -c & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \\ \mathbf{H}_{RL} &= \begin{pmatrix} x & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned} \quad (5.14)$$

Wendet man die so gebildete Homographiematrix auf die Gleichung 2.32 an, so erkennt man, dass das Vorzeichen der Disparität ausschließlich durch das Vorzeichen der Z -Komponente des Objektpunktes beeinflusst wird.

Ein Paar korrespondierender Punkte, die eine Disparität mit alternierendem Vorzeichen aufweisen würden, könnten also nur von einem Objektpunkt P , der hinter dem Stereosystem läge, herrühren. Es ist also offensichtlich, dass die Menge der auftretenden Disparitäten ein einheitliches Vorzeichen ausweisen muss. Die Ausprägung des Vorzeichens ist dann lediglich abhängig von der Wahl des Bezugskoordinatensystems. Hier wurde die linke Kamera als Referenzsystem gewählt. Somit treten ausschließlich positive Werte für die Disparität auf. Alle Punkte P_R^* die zunächst als vermeintliche korrespondierende Punkte erkannt wurden, aber eine negative Disparität aufweisen, werden nun wieder verworfen.

5.4.3 Pointmatching für den Rand

Für die Bildbereiche, welche die Ränder zeigen, stellt sich die Suche nach dem korrespondierenden Punktpaar anders als bei der Lösung des Korrelationsproblems für das Schüttgut dar. Auch hier ist es möglich, die Korrelationsfunktion auf die Datensätze der segmentierten Ränder anzuwenden. Dies würde aber unnötig großen Berechnungsaufwand bedeuten.

Aus der Abbildung 5.15 wird aber deutlich, dass der Schnittpunkt der Epipolarlinie mit den extrahierten Kanten der Ränder die Lösung für das Korrelationsproblem der Ränder darstellt.

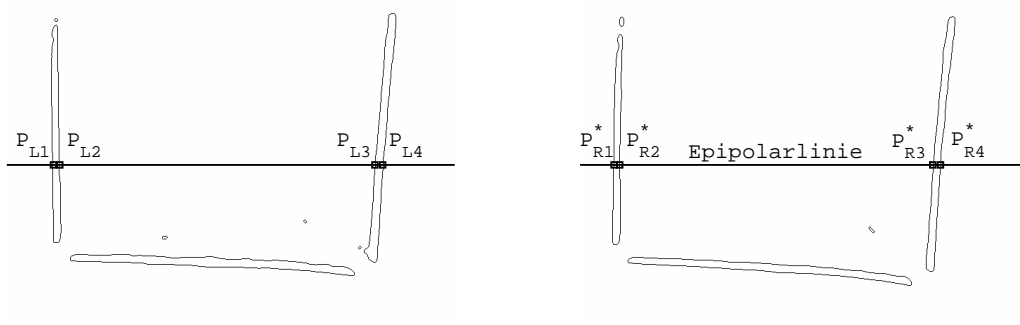


Bild 5.15: Pointmatching für den Rand

5.5 Rekonstruktion

Das Ziel der Stereovisionsanalyse ist es, mit Hilfe der erfassten Bildpaare ein dreidimensionales Modell der abgebildeten Szene zu erstellen. Im folgenden Abschnitt 5.5.1 wird zunächst das allgemeine geometrische Modell der Triangulation zur Rekonstruktion der Objektpunkte erläutert. Der Abschnitt 5.5.2 geht auf die Erweiterung des Triangulationsmodells für die Anwendung auf die Bilddaten, die mit einem realen Kamerasystem erfasst wurden, ein. Abschließend wird in Kapitel 5.5.3 die eingesetzte Triangulationsmethode sowie deren Voraussetzung erläutert.

5.5.1 Die Triangulation

Mit der Kenntniss der Lage der korrespondierenden Punkte sowie der Modellparameter des Stereosystems ist es möglich, durch eine Triangulation den Objektpunkt zu einem Paar korrespondierender Punkte mit Bezug auf das Kamerakoordinatensystem zu rekonstruieren. Die Abbildung 5.16 zeigt das Prinzip der Triangulation.

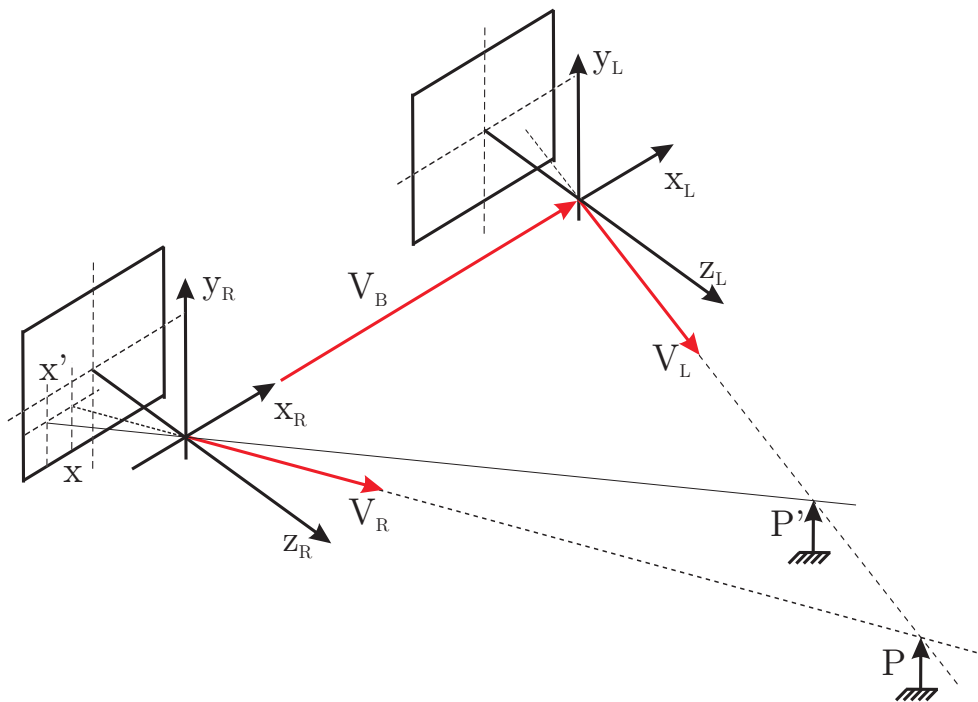


Bild 5.16: Geometrie der Triangulation

Um den Objektpunkt P zu dem gefundenen Paar P_L und P_R korrespondierender Punkte im \mathbb{R}^3 zu rekonstruieren, stellt man das Gleichungssystem des geschlossenen Vektorzugs bestehend aus dem Basisvektor \vec{V}_B und den Abbildungsstrahlen $\lambda\vec{V}_L$ und $\mu\vec{V}_R$ zu den Punkten P_L und P_R auf. Das Gleichungssystem des geschlossenen Vektorzugs ergibt sich nach Gleichung 5.15.

$$\vec{0} = \lambda\vec{V}_L + \mu\vec{V}_R + \vec{V}_B \quad (5.15)$$

Hierbei ergibt sich der Richtungsvektor \vec{V}_i eines Abbildungsstrahls aus dem Ortsvektor zu dem zugehörigen korrespondierenden Punkt P_i . Da der Richtungsvektor der Abbildungsstrahlen immer vom optischen Zentrum der Kamera der Abbildung P_i des Punktes P zeigt, müssen die Werte der Linearfaktoren λ_i und μ_i negativ sein. Analog zu Abbildung 2.3 entspricht die Multiplikation der Richtungsvektoren mit einem negativen Faktor der Punktspiegelung des Vektors am optischen Zentrum der Kamera. Somit zeigt der Richtungsvektor des Abbildungsstrahls vom optischen Zentrum der Kamera zum Objektpunkt P . Wählt man den Ursprung des Bezugskoordinatensystems im optischen Zentrum der linken Kamera, so ergibt sich der Ortsvektor \vec{V}_P zu dem Objektpunkt P durch Einsetzen der Lösung aus Gleichung 5.15 in die Gleichungen 5.16 oder 5.17

$$\vec{V}_{P_i} = \lambda_i\vec{V}_L \quad (5.16)$$

$$\vec{V}_{P_i} = \mu_i\vec{V}_R + \vec{V}_B \quad (5.17)$$

Hierbei liefern die Gleichungen 5.16 und 5.17 gleichwertige Lösungen.

5.5.2 Die erweiterte Triangulation

Die Anwendung der in Kapitel 5.5.1 beschriebenen Methode zur Triangulation der Objektpunkte setzt voraus, dass sich die Abbildungsstrahlen, die durch die korrespondierenden Punkte definiert werden, im dreidimensionalen Raum \mathbb{R}^3 schneiden. Für die ideale Abbildung auf eine kontinuierliche Bildebene ist diese Annahme zulässig. Werden die Bilder mit einer digitalen Kamera erfasst, liegen die Bildpunkte, wie in Kapitel 2.4 beschrieben, diskretisiert vor. Unter dieser Voraussetzung kann nicht mehr davon ausgegangen werden, dass sich die Abbildungsstrahlen, die durch die korrespondierenden Punkte definiert werden, im Objektpunkt schneiden. Nach Emanuele Trucco [10] stellt für diesen Fall die Lösung des Problems

5.18 eine optimale Lösung dar.

$$\vec{0} = \lambda \vec{V}_L + \mu \vec{V}_R + \gamma (\vec{V}_L \times \vec{V}_R) + \vec{V}_B \quad (5.18)$$

Das Kreuzprodukt $\vec{V}_L \times \vec{V}_R$ beschreibt einen Vektor, der senkrecht sowohl auf \vec{V}_L als auch auf \vec{V}_R steht. Dieser Vektor \vec{V}_c definiert somit die kürzeste Verbindungslinie zwischen den windschiefen Abbildungsstrahlen. Die Abbildung 5.17 verdeutlicht die Geometrie der erweiterten Triangulation. Der Ort des so rekonstruierten Punktes ergibt sich dann nach den Gleichungen 5.19 und 5.20

$$\vec{P}_i = \lambda_i \vec{V}_L - \frac{1}{2} \gamma_i \vec{V}_c \quad (5.19)$$

$$\vec{P}_i = \mu_i \vec{V}_R + \vec{V}_B + \frac{1}{2} \gamma_i \vec{V}_c \quad (5.20)$$

Die Gleichungen 5.19 und 5.20 liefern auch in diesem Fall gleichwertige Ergebnisse.

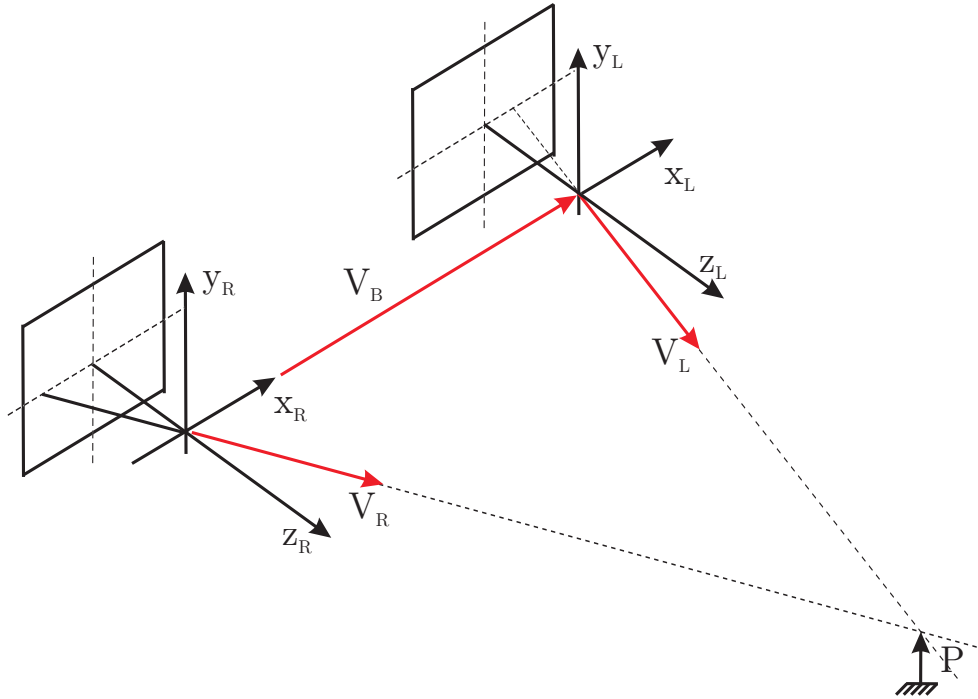


Bild 5.17: Geometrie der erweiterten Triangulation

5.5.3 Triangulation rektifizierter Bildpunkte

Eine besonders einfache Methode für die Triangulation der Objektpunkte ergibt sich, wenn sichergestellt ist, dass sich die rekonstruierten Abbildungsstrahlen S_{Li} und S_{Ri} zu einem be-

liebigen Objektpunkt P_i im Raum schneiden. Diese Annahme ist zulässig, wenn die optischen Zentren O_L und O_R der beteiligten Kameras sowie das Paar korrespondierender Punkte P_L und P_R jeweils auf einer gemeinsamen horizontalen Linie liegen. Für diesen Fall reduziert sich die Triangulation des Objektpunktes $P \in \mathbb{R}^3$ auf ein zweidimensionales Problem. Die Rektifizierung nach Kapitel 3.5.1 transformiert die Bilder, sodass sie einer idealen Abbildung in zwei identischen, koplanaren und achsparallelen Kameras hinreichend nahe kommen. Damit gilt für die Y - und Z -Koordinaten der optischen Zentren O_L und O_R der beteiligten Kameras:

$$\begin{aligned}\vec{O}_{Ly} &= \vec{O}_{Ry} \\ \vec{O}_{Lz} &= \vec{O}_{Rz}\end{aligned}\tag{5.21}$$

Somit ist nach Gleichung 5.10 sichergestellt, dass die Epipolarlinie l_{P_L} im rechten Bild zu einem Punkt P_L im linken Bild eine Horizontale ist.

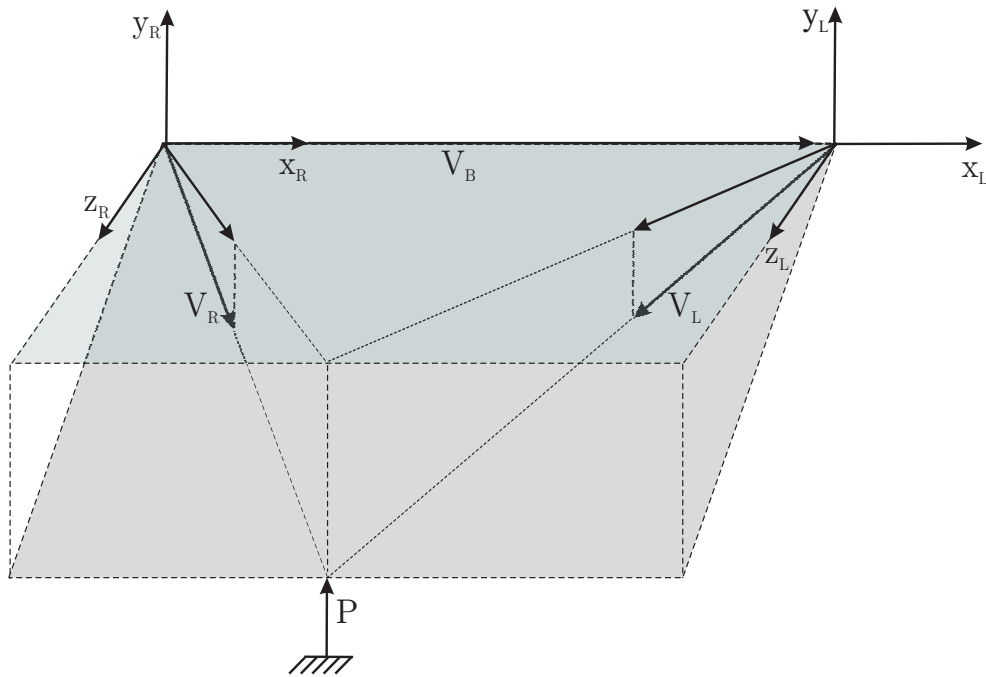


Bild 5.18: Geometrie der Triangulation bei rektifizierten Bildern

Da nach Gleichung 2.31 der korrespondierende Punkt P_R zu einem Punkt P_L ein Element der zugehörigen Epipolarlinie l_{P_L} sein muss, kann angenommen werden, dass die korrespondierenden Punkte immer Element einer gemeinsamen horizontalen Linie sind. Unter dieser Voraussetzung reduziert sich die Rekonstruktion eines Objektpunktes P wieder auf den idealen Fall nach Gleichung 5.15.

$$\vec{0} = \lambda \vec{V}_L + \mu \vec{V}_R + \vec{V}_B$$

Hierbei bezeichnen \vec{V}_L , \vec{V}_R und \vec{V}_B die Vektoren der Abbildungsstrahlen der Punkte P_L und P_R . \vec{V}_B bezeichnet den Basisvektor des Epipolarsystems. Es liegt also ein Problem mit den zwei Unbekannten λ und μ vor. Demnach ist es hinreichend, ein Gleichungssystem mit zwei unabhängigen Gleichungen für die Lösung dieses Problems aufzustellen. Die Abbildung 5.18 verdeutlicht die Geometrie der Triangulation bei einem rektifizierten Bildpaar. Man erkennt, dass es ausreichend ist, das Gleichungssystem unter Verwendung von zwei Komponenten der Vektoren zu lösen. In diesem Fall werden für die Lösung des Problems die X - und Z -Komponenten verwendet. Dies entspricht einer Projektion der Epipolarebene π_P zu dem rekonstruierten Punkt P auf die ZX -Ebene des Epipolarsystems.

6 Ableitung des Füllstandes

In den vorhergehenden Kapiteln 4 und 5 wurde beschrieben, wie aus den eingezogenen Bildern ausgewählte Pixel zunächst die im Bild sichtbaren Objekten zugeordnet und dann durch Triangulation die Objektpunkte im \mathbb{R}^3 rekonstruiert wurden. Aus den so generierten Punktwolken soll nun der Füllstand des Transportmittels abgeleitet werden. Hierfür wird in Abschnitt 6.2 die verwendete Definition der Messgröße *Füllstand* erläutert. Für die Bemessung des Füllstandes muss zunächst ein Koordinatensystem für den Laderaum definiert werden. Der Abschnitt 6.4 erläutert, wie aus den rekonstruierten Punkten, die den Rand repräsentieren das Koordinatensystem für den Laderaum abgeleitet wird. Anschließend müssen die Punkte, die das Ladegut repräsentieren vom Koordinatensystem der Kamera in das Koordinatensystem des Laderaumes transformiert werden. Die Koordinatensystemtransformation wird in Abschnitt 6.5 beschrieben. Abschließend wird in Abschnitt 6.6 darauf eingegangen, wie die ermittelten Daten für die Bemessung des Füllstandes interpretiert werden.

6.1 Der Prozess zur Ableitung des Füllstandes

Die im vorhergehenden Kapitel 5 ermittelten, Objektpunkte werden im Sensorkoordinatensystem angegeben. Da es aber sinnvoller ist, die lokale Füllhöhe des Ladegutes auf das Koordinatensystem des Laderaumes zu beziehen, wird in diesem Funktionsblock das Koordinatensystem des Laderaumes bestimmt und die Messpunkte, die das Ladegut repräsentieren, in dieses Koordinatensystem transformiert. Die Details zur Ableitung des Füllstandes bezüglich des Koordinatensystems des Laderaumes werden in Abschnitt 6.4 beschrieben. Basierend auf den Messpunkten, die den Rand repräsentieren, wird zunächst über eine lineare Ausgleichsrechnung die Ebene bestimmt, welche der obere Rand des Laderaumes definiert. Diese Ebene definiert das im Weiteren verwendete Koordinatensystem des Laderaumes bezüglich der Weltkoordinaten. Anschließend werden die Punkte, die das Ladegut repräsentieren, vom Weltkoordinatensystem in das Koordinatensystem des Laderaumes transformiert. Aufgrund der Auswahl der Messpunkte anhand ihres Informationsgehaltes liegen die Messpunkte nicht uniform verteilt vor. Um nun an jedem beliebigen Punkt im Bereich des Laderaumes die

Füllhöhe angeben zu können, wird eine *konvexe Hülle* mit Hilfe des *XPand*-Algorithmus über die Messpunkte des Ladegutes berechnet. Das Ergebnis ist eine Matrix, die über einem strukturiertem Gitter der *XY*- Ebene des Laderaumes, die Messwerte der lokalen Füllhöhe in Millimetern angegeben, enthält.

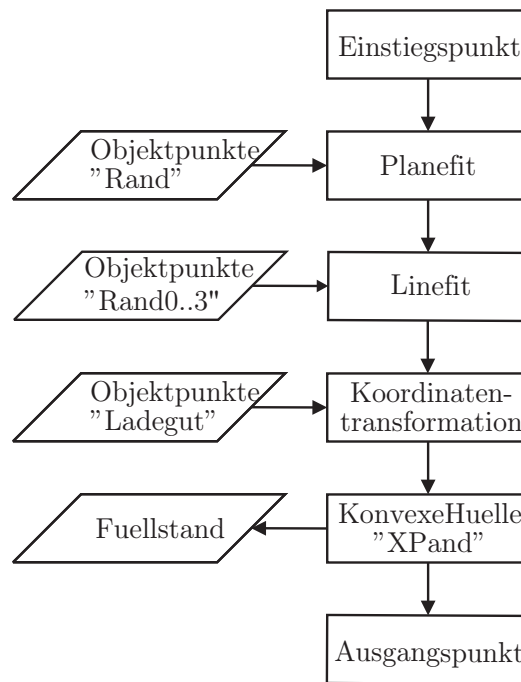
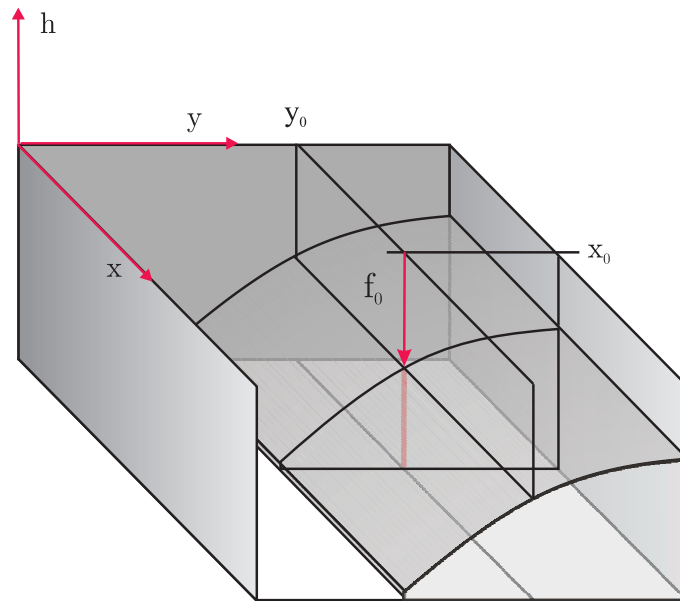


Bild 6.1: Flussdiagramm der Stereovisionsanalyse des Ladegutes

6.2 Die Definition der Messgröße Füllstand

Für die betrachteten Schüttgüter kann nicht angenommen werden, dass sie sich im Laderaum verteilen und diesen wie z.B. ein Fluid ausfüllen. Ohne eine Änderung des Beladungspunktes ist eine vollständige Ausnutzung des Transportvolumens im Allgemeinen nicht erreichbar. Darüber hinaus treten bei gehäckseltem Mais und Gras starke Schwankungen der Materialeigenschaft bezüglich ihres Rieselverhaltens auf. Die Ursache für dieses Verhalten ist unter anderem auf den variierenden Feuchtigkeits- und Stärkegehalt des Schüttgutes zurückzuführen. Somit ist die Formulierung eines mathematischen Modells zur Beschreibung des Rieselverhaltens ohne die Kenntniss der momentanten Materialeigenschaften nur schwer möglich. Für eine numerische Erfassung der Verteilung des Ladegutes im Laderaum ist es sinnvoll, an hinreichend vielen Stellen im Laderaum die lokale Füllhöhe zu erfassen, um somit ein numerisches dreidimensionales Modell der Ladegutoberfläche zu erhalten. Es ist naheliegend

die Objektpunkte $\overrightarrow{P_{i,G}}$ des Objektes *Ladegut* auf ein ortsfest mit dem Laderaum verbundenes Koordinatensystem zu beziehen. Geht man zunächst davon aus, dass ein Gefäß als vollständig gefüllt angesehen werden kann, wenn der Gutspegel den oberen Rand des Gefäßes erreicht, so ist es auch naheliegend, die rekonstruierten Objektpunkte $\overrightarrow{P_{i,G}}$, die dem Ladegut zugeordnet wurden, auf den oberen Rand des Laderaumes zu beziehen. An dieser Stelle sei angemerkt, dass die Betrachtung von Schüttgutmodellen, die eine Befüllung über den oberen Rand des Laderaumes hinaus berücksichtigen, in dieser Arbeit nicht weiter in die Betrachtung des Füllstandes einbezogen werden. Die rekonstruierten Randpunkte $\overrightarrow{P_{i,B}}$ definieren demnach eine Ebene F_0 über dem Laderaum, die eine Füllhöhe definiert, welche wiederum für jede Stelle im Laderaum den Füllstand $h_{xy} = 100\%$ bestimmt. Allein durch Erfassung der Schüttgutoberfläche und der Ränder des Laderaumes ist es aber im Allgemeinen nicht möglich, die lokale Füllhöhe im Laderaum zu erkennen. Dies würde voraussetzen, dass entweder die Geometrie des Laderaumes in irgend einer Form bekannt wäre oder die Tiefe des Laderaumes gemessen würde. In diesem Fall wird allerdings die Geometrie des Laderaumes weder apriori dem System zur Verfügung gestellt, noch kann davon ausgegangen werden, dass das Messsystem die Geometrie des Laderaumes zu Beginn des Befüllungsvorganges vermisst. Für die Steuerung des Beladungsprozesses enthalten jedoch die *inverse lokale Füllhöhe* und die *lokale Füllhöhe* gleichwertige Information. Für die Bestimmung der *inversen lokalen Füllhöhe* wird jedoch nur die Lage der Schüttgutoberfläche in Bezug auf den oberen Rand des Laderaumes benötigt. Die Abbildung 6.2 veranschaulicht die Definition der Messgröße *Füllstand* durch Angabe der inversen lokalen Füllhöhe.

Bild 6.2: Definition der Messgröße *Füllstand*

Es wird deutlich, dass der lokale Füllstand f_{xy} an der Stelle x,y des Laderaumes nicht durch die momentane lokale Füllhöhe h_{xy} sondern die Länge des Lotvektors von der Ebene F_0 auf die Schüttgutoberfläche bemessen wird. Erfasst man nun an $n \times m$ -Stellen eines rektilinearen Rasters des Laderaumes die inverse Füllhöhe, so kann der Beladungszustand durch die $n \times m$ -Matrix F angegeben werden. Hierbei ergibt sich die Auflösung Δx , Δy des Rasters, für das der Füllstand angegeben wird, nach den Gleichungen 6.1 und 6.2.

$$\Delta x = \frac{(x_{max} - x_{min})}{n} \quad (6.1)$$

$$\Delta y = \frac{(y_{max} - y_{min})}{m} \quad (6.2)$$

6.3 Die Bestimmung der Ebene F_0

Mit Hilfe der rekonstruierten Objektpunkt \vec{P}_B , die dem Objekt *Rand* zugeordnet werden können, wird die Ebene F_0 , die durch den oberen Rand des Laderaumes definiert wird, bestimmt. Auf Grund von Fehlern der Lösung des Korrelationsproblems sowie Diskretisierungsfehlern bei der Abbildung der Ränder auf den Kamerasensor muss davon ausgegangen werden, dass die Objektpunkte nicht exakt in der F_0 -Ebene liegen. Aus diesem Grund wird die F_0 -Ebene durch eine lineare Ausgleichsrechnung optimal an die Menge der Randpunkte angenähert. Eine allgemeine Ebene E im \mathbb{R}^3 kann nach Gleichung 6.3 durch die Hessesche Normalform beschrieben werden.

$$E := k = n_1x_1 + n_2x_2 + n_3x_3 \quad (6.3)$$

$$k = \textit{konstant}$$

Die Koeffizienten n_1 , n_2 und n_3 geben hierbei die Linearfaktoren der Einheitsvektoren \vec{x}_1 , \vec{x}_2 und \vec{x}_3 entlang der X -, Y - und Z -Achse des betrachteten Koordinatensystems an. Der Wert k bezeichnet die Länge des Normalenvektors \vec{n}_E vom Ursprung auf die Ebene E und bezeichnet somit den kleinsten Abstand der Ebene vom Ursprung. Stellt man nun die Gleichung 6.3 nach n_3x_3 um, so kann bei bekannten Koeffizienten n_i mit $i = 1 \dots 3$ die x_3 -Komponente zu jedem beliebigen Punkt, der ein Element der Ebene E ist, berechnet werden. Die Gleichung 6.3 wird vektoriell in der Form 6.4 dargestellt.

$$(n_1 \ n_2 \ n_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \quad (6.4)$$

Für ein überbestimmtes Gleichungssystem der Ebene E lässt sich die Gleichung 6.4 in der Form 6.5 formulieren.

$$[\vec{n}_1 \ \vec{n}_2 \ \vec{n}_3] \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{k} \quad (6.5)$$

$$\text{bzw.} \quad \mathbf{P_E} \vec{X} = \vec{k}$$

Hierbei enthalten die Vektoren $\vec{n}_1, \vec{n}_2, \vec{n}_3$ sowie \vec{k} , die Parameter der Punkte, die ein Element der Ebene E sind. Die Vektoren \vec{n}_1, \vec{n}_2 und \vec{n}_3 werden in der Matrix $\mathbf{P_E}$, welche die Steigung der Ebene E entlang der X -, bzw. Y -Achse beschreibt. Für ein überbestimmtes Gleichungssystem gilt Gleichung 6.6

$$\mathbf{P_E} \in \mathbb{R}^{3 \times n} \quad (6.6)$$

$$n > 3$$

Für eine gegebene Punktemenge lässt sich nun durch Lösung des Minimalproblems 6.7 eine Ebene E , die bezüglich des absoluten quadratischen Fehlers optimal an die Punktemenge angenähert ist.

$$Err(x_0) = |\mathbf{P_E} \vec{X} - \vec{k}|^2$$

$$Err(x) \stackrel{!}{=} \min \quad (6.7)$$

$$\Rightarrow \text{grad}(Err(x)) \stackrel{!}{=} 0$$

$$\mathbf{P_E}^T \mathbf{P_E} \vec{X} = \mathbf{P_E}^T \vec{k} \quad (6.8)$$

Bezeichnet man die $3 \times n$ -Matrix, welche die Menge der n -Randpunkte im \mathbb{R}^3 enthält mit $\mathbf{P_B}$, so ergibt sich für diesen Fall die Matrix $\mathbf{P_{F_0}}$, der Ebene F_0 , aus den Spaltenvektoren der Matrix $\mathbf{P_B}$, welche die X - und Y -Komponenten der Randpunkte enthalten. Die dritte Spalte der Matrix $\mathbf{P_{F_0}}$ wird mit Einsen aufgefüllt. Der Vektor $\vec{k_{F_0}}$ enthält die Z -Komponenten der Randpunktematrix $\mathbf{P_B}$, $\vec{P_{B,z}}$. Somit ergibt sich die Ebene F_0 , die optimal an die rekonstruierten Punkte des Laderaumrandes angenähert ist, durch Lösung der Gleichung 6.8.

$$\mathbf{P_{F_0}}^T \mathbf{P_{F_0}} \vec{X} = \mathbf{P_{F_0}}^T \vec{k_{F_0}} \quad (6.9)$$

$$\text{mit} \quad \mathbf{P_{F_0}} = [\vec{P_{B,x}} \ \vec{P_{B,y}} \ \vec{1}]$$

$$\text{und} \quad \vec{k_{F_0}} = \vec{P_{B,z}}$$

Der Lösungsvektor $\overrightarrow{X_{F_0}}$ enthält dann die Steigung der Ebene F_0 in Richtung der aufsteigenden X - und Y -Achse sowie die Länge des Aufpunktvektors.

$$\overrightarrow{X_{F_0}} = \begin{pmatrix} x_{1,F_0} \\ x_{2,F_0} \\ x_{3,F_0} \end{pmatrix} \quad (6.10)$$

Nachteilig an diesem Verfahren ist, dass auch die Punkte, die aufgrund von Messfehlern deutlich außerhalb der Ebene des Randes rekonstruiert wurden, die Lage mit beeinflussen. Um diesem Effekt entgegenzuwirken, werden iterativ die Punkte, deren Abstand $d_{P_{Bi}}$ größer ist als der Plausibilisierungswert d_{krit} eliminiert. Mit der Menge der verbleibenden Punkte wird dann die F_0 -Ebene erneut ermittelt. Dieses Verfahren wird solange ausgeführt, bis sich die Ebenenparameter nicht mehr ändern.

Der Abstand eines Punktes zur Ebene F_0 ergibt sich nach Gleichung 6.11. Nach der Vorschrift 6.15 werden anschließend diejenigen Punkte, deren Abstand zur Ebene F_0 kleiner ist als der Plausibilisierungswert d_{krit} , für eine erneute Bestimmung der Ebene F_0 verwendet.

$$d_{P_{Bi}} = \frac{|\overrightarrow{A_{F_0}P_{Bi}} \cdot \overrightarrow{n_{F_0}}|}{|\overrightarrow{n_{F_0}}|} \quad (6.11)$$

In Gleichung 6.11 bezeichnet $\overrightarrow{A_{F_0}P_{Bi}}$ den Differenzvektor vom Aufpunkt $\overrightarrow{A_{F_0}}$ auf die Ebene zum betrachteten Punkt $\overrightarrow{P_{Bi}}$ und $\overrightarrow{n_{F_0}}$ den Normaleninheitsvektor auf die Ebene F_0 . Die Vektoren $\overrightarrow{A_{F_0}}$ und $\overrightarrow{n_{F_0}}$ haben die gleiche Richtung und ergeben sich direkt aus der Ebene F_0 nach Gleichung 6.12, 6.13 und 6.14.

$$\overrightarrow{V_{F_0}} = \begin{pmatrix} 0 \\ 0 \\ x_{3,F_0} \end{pmatrix} \quad (6.12)$$

$$\overrightarrow{n_{F_0}} = \begin{pmatrix} -x_{1,F_0} \\ -x_{2,F_0} \\ 1 \end{pmatrix} \quad (6.13)$$

$$\overrightarrow{n_{F_0}} = \frac{\overrightarrow{n_{F_0}}}{|\overrightarrow{n_{F_0}}|}$$

$$\overrightarrow{A_{F_0}} = \overrightarrow{V_{F_0}} \cdot \overrightarrow{n_{F_0}} \quad (6.14)$$

Der Vektor \vec{A}_{F_0} ergibt sich aus Gleichung 6.14 als Projektion eines beliebigen Vektors (hier \vec{V}_{F_0}) zu einem Punkte P_{F_0} der Ebene auf den Vektor \vec{n}_{F_0} . Die Menge der weiterverwendeten Punkte ergibt sich dann nach:

$$\begin{aligned} \mathbf{P}_B^* &:= \mathbf{P}_B \\ &\text{mit} \\ d_{krit} &\geq d(P_{Bi}) \end{aligned} \tag{6.15}$$

Für den Plausibilisierungswert hat sich der Wert $d_{krit} = 30cm$ als geeignet gezeigt. Für den hier vorliegenden Abstand des Messobjektes vom Kamerasystem entspricht dieser Wert dem mittleren Positionsfehler eines rekonstruierten Punktes, wenn für die Lösung des Korrelationsproblems ein Fehler von einem Pixel angenommen wird.

Die folgende Tabelle 6.1 zeigt die Parameter der F_0 -Ebene der ersten drei Iterationsschritte für die Auswertung eines repräsentativen Bildpaares.

Iterationsschritt	x_{1,F_0}	x_{2,F_0}	x_{3,F_0}	$var(d_{P_{Bi}})$
1	0.2911	0.1096	1005.3	522.84
2	0.2898	0.1129	999.1	194.76
3	0.2898	0.1129	999.1	194.76

Tabelle 6.1: Ebenenparameter

Es hat sich gezeigt, dass schon nach der ersten Iteration nur noch eine unwesentliche Verbesserung der Genauigkeit in der Definition der F_0 -Ebene erreicht wird. Somit wird der Algorithmus bereits nach der zweiten Iteration abgebrochen. Eine weitere Verbesserung der Genauigkeit in der Bestimmung der F_0 -Ebenen wäre nur durch den Einsatz von Kameras mit höherer Auflösung oder anderer Auswertungsalgorithmen für das Korrelationsproblem möglich.

6.4 Koordinatensystem des Laderaumes

Während der Messung befinden sich die beteiligten Fahrzeuge in relativer Bewegung zueinander. Da die rekonstruierten Punkte zunächst auf das Kamerakoordinatensystem bezogen berechnet werden, hat das zur Folge, dass die rekonstruierten Punkte bei jeder Messung an einem neuen Ort in diesem Koordinatensystem erscheinen. Eine Füllstandsermittlung mit Bezug auf das Kamerakoordinatensystem wäre also umständlich und nur wenig sinnvoll.

Es ist aber naheliegend, den Füllstand auf ein Koordinatensystem zu beziehen, das ortsfest mit dem zu befüllenden Laderaum verbunden ist. Dies setzt die Identifikation dieses Koordinatensystems voraus.

Wie in Abbildung 6.2 dargestellt, soll das Koordinatensystem des Laderaumes anhand seiner Ränder definiert werden. Hierfür muss die Lage der Vektoren bestimmt werden, die jeweils einen sichtbaren Rand des Laderaumes repräsentieren. Da bereits im vorhergehenden Abschnitt 6.3 die Ebene F_0 , in der die Ränder liegen, identifiziert wurde, ist die Lage der Z -Achse des Koordinatensystems des Laderaumes bereits bestimmt worden. Das Problem reduziert sich also auf das zweidimensionale Problem der Identifikation der Lage der X - und Y -Achse in der Ebene F_0 . Aufgrund des im Allgemeinen hinreichend rechteckigen Grundrisses des Laderaumes kann zudem die Restriktion eingeführt werden, dass die X - und Y -Achse des Laderaumes rechtwinklig zueinander stehen und zusammen mit der Z -Achse ein Rechtssystem aufspannen müssen. Dieser Sachverhalt wird wiederum durch eine lineare Ausgleichrechnung approximiert. Die Abbildung 6.3 zeigt den rekonstruierten Rand des Laderaumes sowie das identifizierte Koordinatensystem des Laderaumes im Kamerakoordinatensystem.

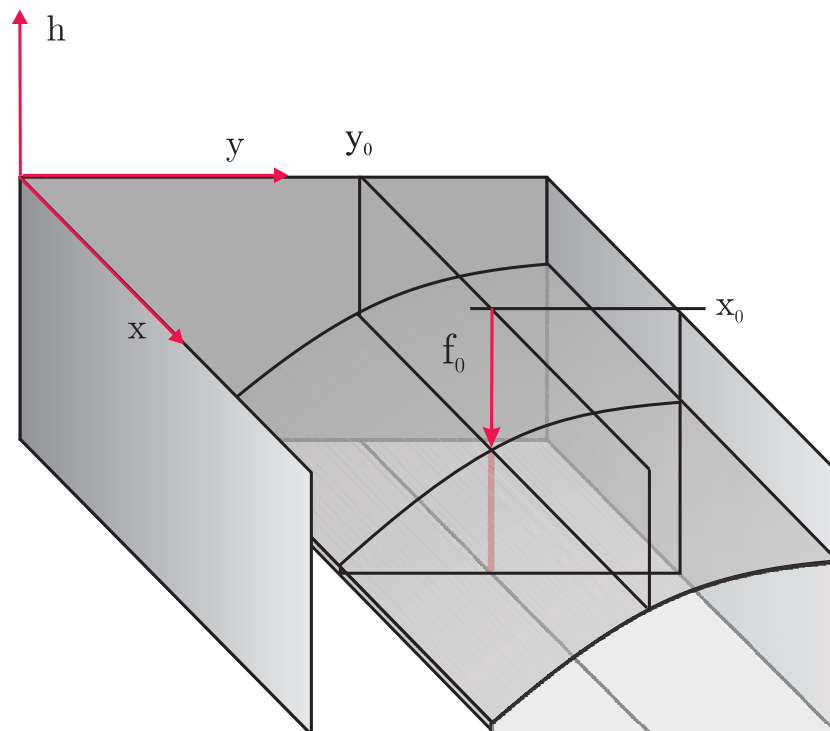


Bild 6.3: Koordinatensystem des Laderaumes

Bezeichnet man die Punkte des sichtbaren Randes mit R , so kann man den Vektor, der die Untermenge der Punkte, die eine der vier Seiten des Randes repräsentieren, mit \vec{R}_i bezeichnen. Die Parameterdarstellung einer allgemeinen Geraden im \mathbb{R}^2 lautet:

$$g : c = ax + by \quad (6.16)$$

Somit lässt sich für jeden Vektor \vec{R}_i ein überbestimmtes Gleichungssystem zur Bestimmung der Geraden g_{R_i} , die den i -ten Rand repräsentiert, ausstellen.

$$\vec{c} = \vec{X}\mathbf{R} \quad (6.17)$$

Eine Gerade g_{R_i} im \mathbb{R}^2 , die eine Seite des Randes repräsentiert, kann dann durch Lösen des Problems 6.8 gefunden werden.

6.5 Transformation der Koordinatensysteme

Die rekonstruierten Objektpunkte der Objekte *Rand* und *Schüttgut* sind als Ergebnis der Triangulation (vgl. Abschnitt 5.5.1) im Koordinatensystem S_K des Epipolarsystems angegeben. Das Ziel besteht darin, die Objektpunkte, die die Schüttgutoberfläche repräsentieren, in das Koordinatensystem des Laderaumes zu transformieren. Hierfür müssen alle Punkte der Schüttgutoberfläche mit den Vektoren \vec{R}_R und \vec{t}_R transformiert werden, die den Versatz zwischen dem Kamerakoordinatensystem und dem Koordinatensystem des Laderaumes beschreiben. Der Vektor \vec{R}_R beschreibt an dieser Stelle den Winkelversatz und der Vektor \vec{t}_R den translativen Versatz der Koordinatensysteme. Entsprechend der Gleichung 2.1 bildet die Gleichung 6.18 einen Objektpunkt der rekonstruierten Schüttgutoberfläche im Koordinatensystem des Laderaumes ab.

$$P_G|_{S_R} = [\mathbf{R}_R \vec{t}_R] P_G \quad (6.18)$$

Die Rotationsmatrix \mathbf{R}_R setzt sich hierbei wieder gemäß Gleichung 2.3 aus der Kombination der Matrizen für die Einzeldrehungen um die Achsen des Kamerakoordinatensystems zusammen. Mit Hinsicht auf den Berechnungsaufwand ist es sinnvoll, die Rotation entsprechend der Vorschrift nach Rodrigues [51] auszuführen. Die Vorschrift nach Rodrigues beschreibt eine Rotation im dreidimensionalen Raum durch eine einzige Rotation um eine Rotationsachse im Raum, somit ist nur eine Matrixoperation pro Punkt nötig. Hierfür muss zunächst der Vektor \vec{R} , der die Rotationsachse beschreibt und der Winkel φ_{nF_0} , die Drehung um die Rotationsachse angibt, bestimmt werden.

Der Vektor \vec{R} ergibt sich aus dem Kreuzprodukt des Normalen-Aufpunktsvektors \vec{A}_{F_0} auf die Ebene F_0 und der Z -Achse. Der Normalen-Aufpunktvektor \vec{A}_{F_0} ergibt sich seinerseits wieder aus dem Vektorprodukt eines Ortsvektors \vec{P}_{F_0} zu einem beliebigen Punkt P_{F_0} auf der Ebene F_0 mit dem Normaleneinheitsvektor \vec{n}_{F_0} der Ebene F_0 nach Gleichung 6.14.

$$\vec{A}_{F_0} = \vec{V}_{F_0} \cdot \vec{n}_{F_0}$$

Die Rotationsachse ergibt sich dann nach Gleichung 6.19.

$$\begin{aligned} \vec{V}_{RF_0} &= \vec{A}_{F_0} \times \vec{Z}_K \\ \text{mit} \\ \vec{Z}_K &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned} \quad (6.19)$$

Der Versatzwinkel $\varphi_{n_{F_0}}$ um diese Achse ergibt sich dann nach Gleichung 6.20

$$\varphi_{n_{F_0}} = -\arccos\left(\frac{|\vec{Z}_K \cdot \vec{n}_{F_0}|}{|\vec{Z}_K| |\vec{n}_{F_0}|}\right) \quad (6.20)$$

Der Vektor \vec{V}_{RF_0} , der die Rotationsachse beschreibt, kann entsprechend Gleichung 2.39 für die Anwendung des Kreuzproduktes als Matrix $\mathbf{V}_{\mathbf{RF}_0}$ dargestellt werden. Die Matrix $\mathbf{V}_{\mathbf{RF}_0}$ ergibt sich dann nach Gleichung 6.21.

$$\mathbf{V}_{\mathbf{RF}_0} = \begin{pmatrix} 0 & -V_{RF_0,z} & V_{RF_0,y} \\ V_{RF_0,z} & 0 & -V_{RF_0,x} \\ -V_{RF_0,y} & V_{RF_0,x} & 0 \end{pmatrix}_{\times} \quad (6.21)$$

Somit ergibt sich die *Rodrigues*-Rotationsmatrix nach Gleichung 6.22

$$\mathbf{R}_{\mathbf{Rod}} = \mathbf{I} + \sin(\varphi_{n_{F_0}}) \mathbf{V}_{\mathbf{RF}_0} + (1 - \cos(\varphi_{n_{F_0}})) \mathbf{V}_{\mathbf{RF}_0}^2 \quad (6.22)$$

Mit Gleichung 6.22 ergibt sich die Transformation eines Punktes \vec{P}_G der Schüttgutoberfläche aus dem Koordinatensystem S_K des Epipolarsystems in das Koordinatensystem S_R des Laderaumes nach Gleichung 6.23

$$\vec{P}_G|_{S_L} = \mathbf{R}_{\mathbf{Rod}} \vec{P}_G|_{S_K} + \vec{t}_R \quad (6.23)$$

Der Translationsvektor \vec{t}_R ergibt sich dann als der negative Normalen-Aufpunktvektor auf die Ebene F_0 .

6.6 Resample

In Vorbereitung für die Suche nach den korrespondierenden Punkten wurden aus dem linken Bild des Stereobildpaares Referenzpunkte gewählt, die aufgrund ihrer Eignung in der Bildebene eindeutig lokalisiert werden zu können, als geeignet bewertet wurden. Dies hat zur Folge, dass die Messpunkte des lokalen Füllstandes nicht uniform im Messbereich verteilt sind. Die Abbildung 6.4 zeigt ein Beispiel für Verteilung der Messpunkte bezogen auf das Koordinatensystem des Laderaumes.

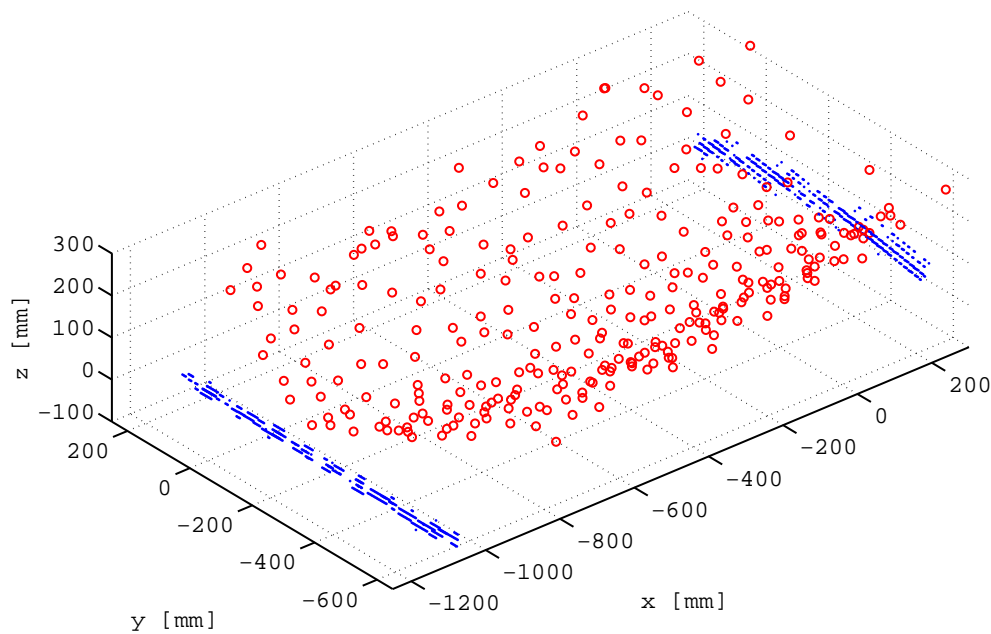


Bild 6.4: Nicht uniform verteilte Messpunkte

In Abbildung 6.4 bezeichnen die roten Kreise die Messwerte des inversen Füllstandes und die blauen Punkte die Lage zweier Ränder des Laderaumes im Koordinatensystem des Laderaumes. Um an jeder beliebigen Stelle des Laderaumes den Füllstand schätzen zu können muss zwischen den Messpunkten interpoliert werden. Um die Interpolation der Messwerte für weitere Anwendungen einfach zu gestalten, ist es sinnvoll, die Messwerte auf ein rektifiziertes Gitter abzubilden.

An dieser Stelle wird der *Xpand* Algorithmus [52] angewendet. Hierbei wird ein kreisrundes Flächenelement mit einem begrenzten Radius r um die Stelle (x_r, y_r) des zu bestimmenden Messpunktes H_r definiert und in acht gleichgroße Kreissegmente unterteilt. In diesen Kreissegmenten wird jeweils der Messpunkt gesucht, dessen x - und y -Koordinate der Stelle des zu bestimmenden Wertes am nächsten ist. Die Abbildung 6.5 zeigt die Auswahl der zu berücksichtigenden Punkte.

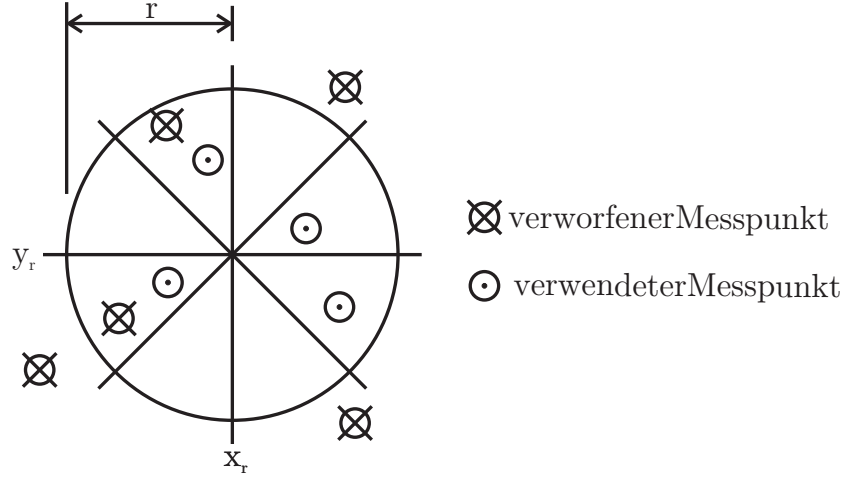


Bild 6.5: XPand - Auswahl der zu verarbeitenden Punkte

Das Zentrum des Kreises definiert die Stelle des zu bestimmenden Messwertes. Die gekreuzten Stellen außerhalb des Radius r werden nicht berücksichtigt, da sie nicht im Einzugsgebiet der Stützstelle liegen. Die gekreuzten Stellen innerhalb des Radius r werden nicht berücksichtigt, da sie im jeweiligen Kreissegment nicht die nächsten Nachbarn zur Stelle des gesuchten Messwertes sind. Somit ergibt sich der zu bestimmende Messwert durch die Bildung des Mittelwertes \bar{h}_{xy} der verbleibenden, hier mit einem Punkt markierten, Messwerte. Der Messwert H_r an einer Stelle x_r, y_r ergibt sich dann nach Gleichung 6.24.

$$H_r = \bar{h}_{x_r y_r} \quad (6.24)$$

$$\bar{h}_{x_r y_r} = \frac{1}{n} \sum_{xy} h_{ij}$$

mit $r > \sqrt{(x_r - x)^2 + (y_r - y)^2}$

Darüber hinaus hat diese Methode noch den positiven Effekt einer Mittelwert- bzw. Tiefpassfilterung auf den Datensatz der rekonstruierten Objektpunkte. Diese Tiefpassfilterung hat zur Folge, dass die Unstetigkeiten im rekonstruierten Datensatz und eventuelle Ausreißer reduziert werden. Die Wirkung des Tiefpassfilters kann durch den Radius r , der die Menge

\mathbb{P} zu berücksichtigender Punkte definiert, eingestellt werden. Es ist an dieser Stelle zwingend notwendig, den Radius r deutlich größer als den in Kapitel 5.2 gewählten Mindestabstand zwischen den markierten Bildpunkten zu wählen, da sonst in den Umgebungen um die Stützstellen keine weiteren Stützstellen gefunden werden können. Durch die Definition des Wertebereichs der XY -Ebene, über der die Messwerte durch den XPand Algorithmus reorganisiert werden, wird sichergestellt, dass nur Messpunkte, die sich im Bereich der Ladefläche befinden ausgegeben werden. Die Abbildung 6.6 zeigt den inversen Füllstand aus Abbildung 6.4 nach der Anwendung des Resampling auf die Daten bezüglich Koordinatensystem des Laderaumes.

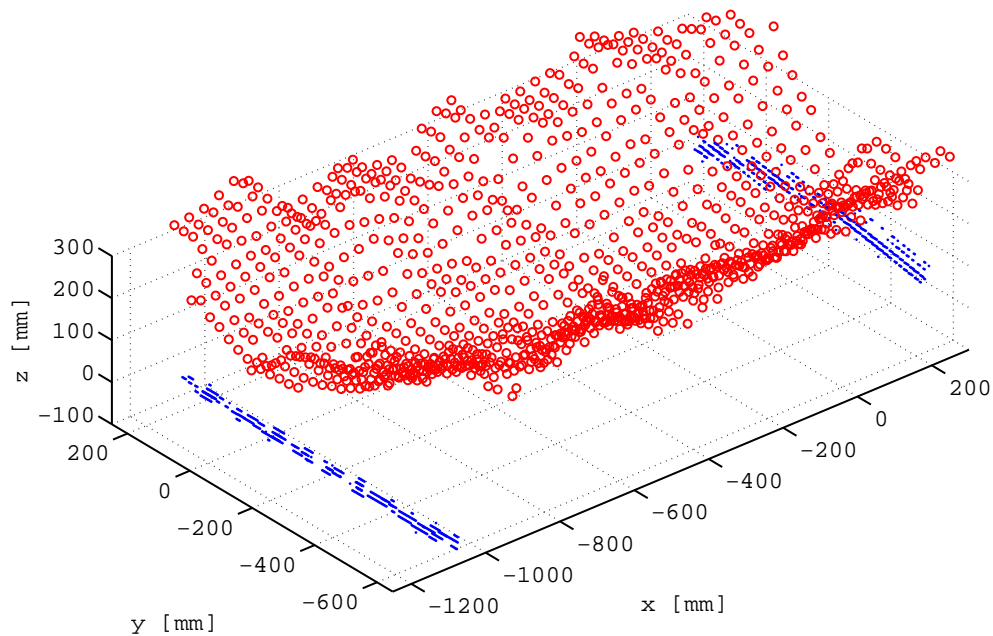


Bild 6.6: Rektilinear verteilte Messpunkte

7 Das Messsystem

Im abschließenden Kapitel wird die Genauigkeit des Messsystems *FILLED* untersucht. Hierfür soll einerseits die Genauigkeit in der Erfassung der Oberfläche und andererseits die Messverzögerung betrachtet werden. An dieser Stelle sei angemerkt, dass die Messverzögerung maßgeblich durch die Leistungsfähigkeit der eingesetzten Computer beeinflusst wird. Mit Hinsicht auf die stetig steigende Berechnungsleistung handelsüblicher Computer ist die Angabe der Messverzögerung auf die zum Zeitpunkt der Entwicklung verfügbare Computertechnik bezogen. Bereits während der ca. dreijährigen Entwicklungszeit des Messsystems *FILLED* ist die Leistung der handelsüblichen Computer von durchschnittlich 2,5 GHz bei 32 Bit Wortbreite auf weit über 3,5 GHz bei 64 Bit Wortbreite angestiegen. Vor diesem Hintergrund werden in absehbarer Zeit Computer zur Verfügung stehen, welche die Messung mit *FILLED* mit deutlich geringerer Messverzögerung ermöglichen.

Das System wird zurzeit mit einem handelsüblichen PC betrieben. Das Referenzsystem hat die folgenden Leistungsdaten:

- Prozessor: Intel Pentium®4, 2,4 GHz, 32Bit
- Speicher: 1024 MByte
- Betriebssystem: SUSE Linux (Kernel 2.6.11)
- Kamerasystem: 2 x IDS Imaging Pacific 9820
- Bildauflösung: 640×480 Pixel, 3 Byte Farbauflösung
- Verwendete Software Bibliotheken: Intel OpenCV Library Ver 1.0, GNU Scientific Library, VTK 4.0

Mit den oben angeführten Spezifikationen beträgt die mittlere Dauer vom Start des Bildinzugs bis zur Bereitstellung der Messdaten im Arbeitsspeicher nicht mehr als $dt = 2,8$ Sekunden. An dieser Stelle sei angemerkt, dass durch Variation der Systemeinstellungen, wie zum Beispiel der Auflösung des DeHough-Raumes oder der Art der Vorsteuerung der DeHough-Transformation (vgl. Kapitel 4.2.5) Messverzögerung sowie die Genauigkeit der

Messergebnisse beeinflusst werden. Die hier angeführten Werte beruhen auf einem Parametersatz, der sich für die Anwendung des Messsystems auf Videodaten aus der Maisernte 2006 als geeignet gezeigt hat. Allgemein kann jedoch die Aussage gemacht werden, dass ein höheres Maß an Genauigkeit auch eine längere Berechnungsdauer zur Folge hat. Im Anhang E sind die wichtigsten Konfigurationdaten des Referenzmesssystems abgelegt.

7.1 Die Fehlerquellen

Das geometrische Modell zur Rekonstruktion der Messpunkte im \mathbb{R}^3 liefert exakte Ergebnisse unter der Voraussetzung, dass ein Paar korrespondierender Punkte tatsächlich die Abbildung von ein und dem selben Objekt ist. Dies kann jedoch mit einem realen Kamerasystem nicht sichergestellt werden. Die wesentlichsten Fehlerquellen hierbei sind:

- Abbildungsfehler der Objektive
- Parameterunsicherheiten der Kameras
- Abbildung auf ein diskretes Gitter
- mangelhafte Übereinstimmung der korrespondierenden Punkte

Das geometrische Modell - die Triangulation - welches für die Rekonstruktion der Objektpunkte auf der Basis der Paare korrespondierender Punkte genutzt wird, geht davon aus, dass die Abbildung eines Objektpunktes durch ein System idealer Lochkameras geschieht (vgl. 2.2). Wie bereits in Kapitel 2.4 ff. erwähnt wurde, unterliegen die eingezogenen Bilder aber Abbildungsfehlern. Diese Abbildungsfehler haben zur Folge, dass die Abbildung eines Objektpunktes von der Abbildung durch eine ideale Kamera abweicht, was wiederum dazu führt, dass die Position des rekonstruierten Objektpunktes nicht der Position des tatsächlichen Objektpunktes entspricht. Die Abbildung 7.1 veranschaulicht die Rekonstruktion eines Objektpunktes auf der Basis verzerrt abgebildeter Bildpunkte. An Abbildung 7.1 wird deutlich, dass besonders die Tiefeninformation des rekonstruierten Punktes P' durch die verzeichnete Abbildung des Punktes P auf der rechten Bildebene beeinflusst wird. Der gleiche Effekt tritt auf, wenn der Ort des optischen Zentrums der Kameras, welche die Ursprünge der Kamerakoordinatensysteme darstellen, nicht exakt bestimmt werden können.

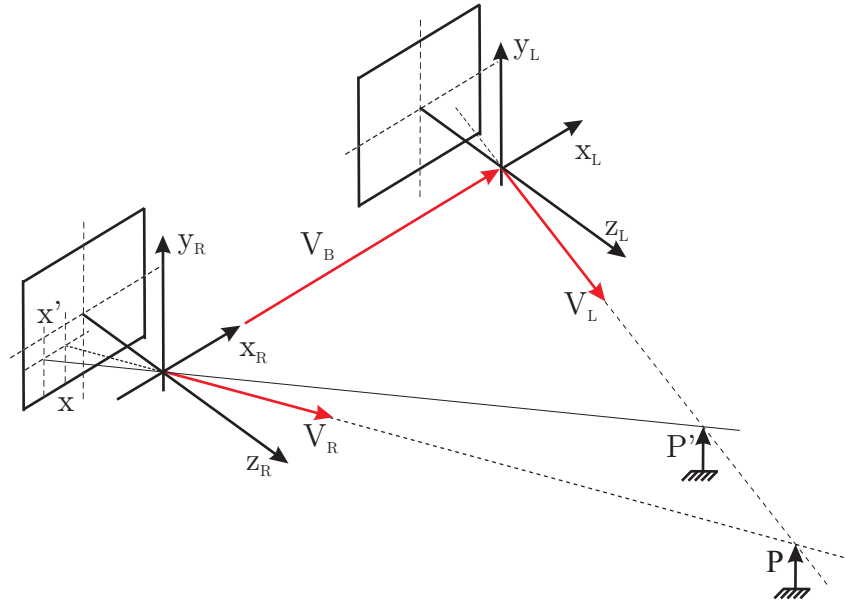


Bild 7.1: Rekonstruktion eines verzeichnet abgebildeten Punktes

7.2 Der Einfluss des Disparitätsfehlers

Wie aus Abbildung 7.1 sichtbar wird, hat eine durch ein fehlerhaftes Matching falsch bestimmte Disparität der korrespondierenden Punkte einen Fehler bei der Rekonstruktion des Objektpunktes zur Folge. Wie in Kapitel 5 beschrieben wurde, kann der Ort des rekonstruierten Punktes P' durch Gleichung 7.1 ermittelt werden.

$$\vec{P'}|_{s_o} = \lambda \vec{V_P} \quad (7.1)$$

Wobei $\vec{V_P}$ den Einheits-Richtungsvektor des Abbildungsstrahls im betrachteten optischen System bezeichnet. Hierbei ergibt sich nach Gleichung 5.22 λ als Funktion der Disparität $D(u_1, u_2)$ des betrachteten Paares korrespondierender Punkte und der Basislänge V_B des stereoptischen Systems.

$$Z = \frac{|V_B|c}{(O_r - u_r) - (O_l - u_l)} \quad (7.2)$$

$$X = \frac{u_l Z}{c}$$

$$Y = \frac{v_l Z}{c}$$

Für die Auswertung digitalisierter Bilder wird deutlich, dass der Einfluss des Matching-Fehlers auf die Position der rekonstruierten Punkte von der Entfernung des Punktes vom optischen System abhängt. Die Positionsabweichung eines rekonstruierten Punktes P' von der Position des abgebildeten Punktes P als Funktion des Matching-Fehlers ergibt sich gemäß Gleichung 7.4 nach

$$\bar{Z} = Z + \Delta Z_{MErr} \quad (7.3)$$

$$Z_i + \Delta Z_{MErr} = \frac{|V_B|c}{(O_r - u_r + \Delta u) - (O_l - u_l)} \quad (7.4)$$

$$\bar{X} = \frac{u_l Z}{c}$$

$$\bar{Y} = \frac{v_l Z}{c}$$

\bar{X} , \bar{Y} und \bar{Z} stehen für die X -, Y - und Z -Koordinate der rekonstruierten Punkte der Schüttgutoberfläche. ΔZ_{MErr} bezeichnet den Fehler der rekonstruierten Z -Koordinate der Schüttgutoberfläche. u_l und u_r bezeichnen die u -Koordinaten der tatsächlich korrespondierenden Punkte im linken bzw. rechten Bild. Der Fehler der u -Koordinate der korrespondierenden Punkte wird Δu genannt. Die Abbildung 7.2 veranschaulicht den Fehler ΔZ der Z -Koordinate des rekonstruierten Punktes als Funktion des Matching-Fehlers Δu und des Abstandes Z des Objektpunktes vom optischen System.

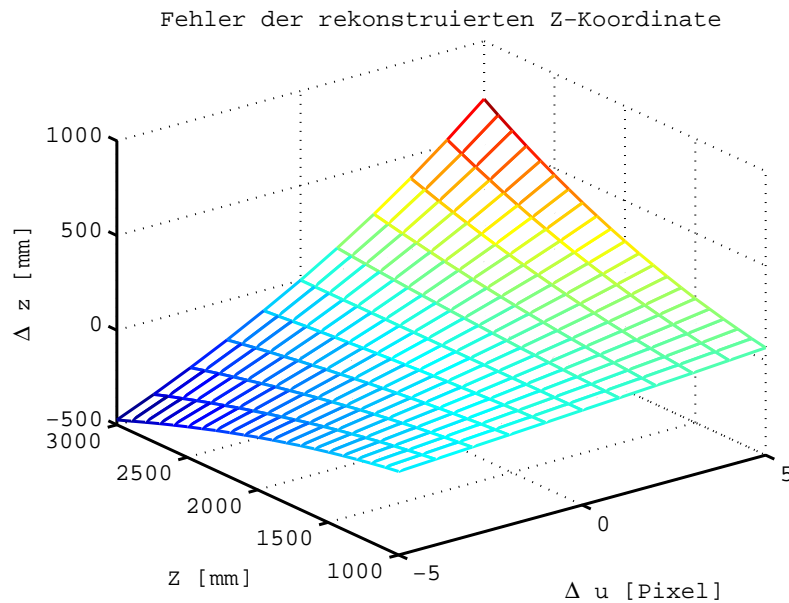


Bild 7.2: Fehler des rekonstruierten Punktes

Bedingt durch die Abbildung des Objektpunkts auf das diskrete Raster des Sensorchips, muss von einem Mindestfehler für das PointMatching von $\Delta u = \pm 1 \text{ Pixel}$ ausgegangen werden. Die Versuche haben gezeigt, dass durch die Auswahl der weiter zu verwendenden Punkte (siehe Kapitel 5.4.1) auf Basis des normierten Korrelationskoeffizienten ein größerer Fehler beim PointMatching nicht auftritt. Bei einer maximalen Entfernung des Messobjektes zum Kamerasystem von $Z_{max} \leq 3000 \text{ mm}$ kann in diesem Prozessschritt von einem absoluten resultierenden Messfehler der Z-Koordinate von $|\Delta Z| < 118 \text{ mm}$ ausgegangen werden. An dieser Stelle sei angemerkt, dass die Z-Koordinate als Linearfaktor für die Bestimmung des Objektpunktes P verwendet wird und somit der Positionsfehler des rekonstruierten Punktes sich proportional zum Fehler der Z-Koordinate verhält.

7.3 Der Einfluss des Abbildungsfehlers

Wie bereits in Abschnitt 3 erwähnt, weicht die Abbildung der betrachteten Szene durch eine reale Kamera von der Abbildung durch eine Lochkamera ab. In Kapitel 3.5.1 wurde bereits beschrieben, wie die eingezogenen Bilder durch Rektifikation an die Abbildung durch ein achsparalleles System idealer Lochkameras angenähert werden. Auch diese Abbildung unterliegt den Parameterunsicherheiten der Kalibrierung, so dass eine exakte Lösung lediglich angenähert werden kann. Die Kompensation der unterschiedlichen Linsenverzeichnung sowie des Versatzes der Kamerakoordinatensysteme hat zur Folge, dass die verschiedenen Bildbereiche unterschiedlich stark transformiert werden. Besonders deutlich wird dieser Fehler bei der Betrachtung der radialen Verzeichnung. In diesem Fall wird ein Bildpunkt auf dem Radius bezüglich des optischen Zentrums der Linse nach außen oder nach innen verzeichnet. Diesen Sachverhalt verdeutlicht die Abbildung 2.8. Die Position des verzeichneten Punktes ergibt sich nach Gleichung 2.10.

$$\tilde{r} = r + k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots$$

Werden nun die Verzerrungskoeffizienten k_1, k_2, \dots, k_n fehlerhaft bestimmt, tritt eine vom Radius des betrachteten Pixels abhängige, fehlerhafte Entzerrung auf. Gleichung 7.5 beschreibt den Fehler des so bestimmten Entzerrungsradius.

$$\tilde{r} + \Delta r = r + (k_1 + \Delta k_1) r^2 + (k_2 + \Delta k_2) r^4 + \dots \quad (7.5)$$

Die durchgezogenen Linien in Abbildung 7.4 stellen die Abbildungsstrahlen der Punkte P_1 und P_2 für die rechte und linke Kamera dar. Die jeweiligen Bildpunkte von P_1 und P_2 werden mit der Entfernung r_{r1} , r_{r2} zum optischen Zentrum in der rechten Kamera abgebildet. In der linken Kamera werden die Bildpunkte mit den Entfernungen r_{l1} bzw. r_{l2} zum optischen Zentrum abgebildet. Die gestrichelten Linien stellen die rekonstruierten und fehlerbehafteten Abbildungsstrahlen der o.a. Punkte dar. Es ist offensichtlich, dass die Punkte \bar{P}_1 und \bar{P}_2 bei zu groß bestimmter Disparität näher am Kamerasystem rekonstruiert werden. Bei zu klein bestimmter Disparität dementsprechend zu weit entfernt. Es wird jedoch auch sichtbar, dass aufgrund des unterschiedlich starken Disparitätsfehlers der Punkt \bar{P}_1 nicht mit dem gleichen Fehler Δz wie der Punkt \bar{P}_2 rekonstruiert wird. Somit wird deutlich, dass durch fehlerhaft bestimmte Verzeichnungskoeffizienten eine Wölbung des rekonstruierten Datensatzes erzeugt wird.

Der Gesamtfehler aus falsch gematchten Punktepaaaren und dem Disparitätsfehler, der aus fehlerhaften Verzerrungskoeffizienten resultiert, ergibt sich durch die Addition der jeweiligen Fehleranteile. Gleichung 7.6 beschreibt den Gesamtfehler der Entfernungsbestimmung.

$$\Delta Z = \frac{|V_0|c}{(O_r - u_r + \Delta u_{r,M} + \Delta u_{r,R}) - (O_l - u_l)} \quad (7.6)$$

Hierbei beschreibt $u_{r,M}$ den Matching-Fehler und $u_{r,R}$ den Fehler, der aus falsch bestimmten Verzerrungskoeffizienten herrührt. An dieser Stelle sei angemerkt, dass bei der Auswertung von rektifizierten Bildern, die Z -Koordinate der rekonstruierten Punkte nur noch von der Disparität D entlang der u -Ache des Bildkoordinatensystems abhängt. Darüber hinaus ist zu bemerken, dass dieser Fehler ausschließlich durch die unterschiedlichen Entfernungen der abgebildeten Punkte zum optischen Zentrum beeinflusst wird. Somit wird der Einfluss dieses Fehlers mit zunehmender Entfernung des abgebildeten und damit auch des rekonstruierten Punktes zum Kamerasystem geringer, da die Winkel der Abbildungsstrahlen für die rechte und linke Kamera zur optischen Achse mit zunehmender Entfernung auf den gleichen Wert konvergieren. Gleiches gilt für die Abbildung bzw. die Rekonstruktion von Punkten, die mittig bezüglich der X -Achse zwischen den Kameras liegen. Auch hier treten die Abbildungsstrahlen im gleichen Winkel zu den optischen Achsen durch die Objektive. Sie unterscheiden sich lediglich durch das Vorzeichen. Das Vorzeichen wird nach Gleichung 7.5 jedoch durch die Wahl eines Polynoms mit geraden Koeffizienten unterdrückt. Dieser Effekt hat die Wölbung des rekonstruierten Datensatzes zur Folge. Die Abbildung 7.5 zeigt einen auf diese Weise verzerrten Datensatz.

Die Objektpunkte, die rekonstruiert wurden, stammen von einer Ebene, die mit dem Stereovisionssystem eingezeichnet wurde, um anschließend rekonstruiert zu werden. Hierbei be-

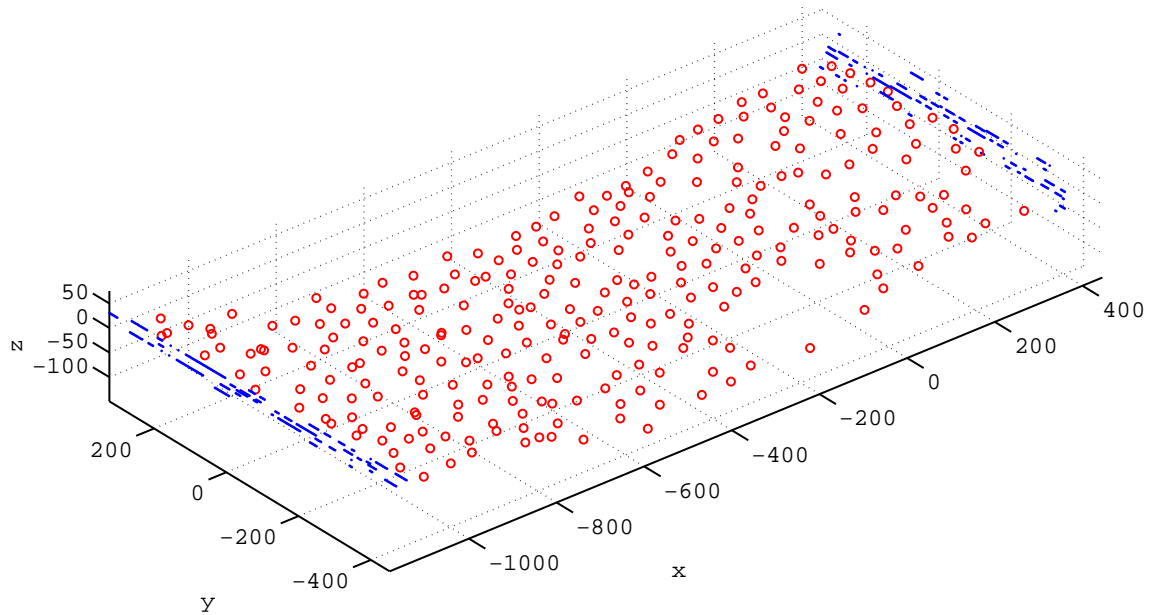


Bild 7.5: Disparitätsfehler durch Rektifizierung

zeichnen die blauen Punkte die rekonstruierten Objektpunkte, die dem Rand zuzuordnen sind, und die roten Punkte, welche dem Ladegut zuzuordnen sind. Es wird deutlich, dass die Punkte des Ladegutes, die um die Koordinaten $x = 0$, $y = 0$ positioniert sind, eine andere Entfernung zum Kamerasystem aufweisen als diejenigen Punkte, die am Rand des Datensatzes gelegen sind. Dieser Effekt ist auf die oben angeführten Ungenauigkeiten in der Rektifikation zurückzuführen. An dieser Stelle sei angemerkt, dass nicht zuletzt mit Hinsicht auf die Anwendung auf einer Landmaschine kostengünstige Kameras und Objektive eingesetzt wurden. Es ist denkbar, dass diese Effekte durch den Einsatz höherwertiger Kameras und Objektive reduziert werden können. Der maximale Fehler, der aufgrund fehlerhafter Verzerrungskoeffizienten auftritt, beläuft sich bei den hier eingesetzten Kameras und den verwendeten Objektiven mit einer Brennweite von $f_c = 6 \text{ mm}$ auf $\Delta Z_R < \pm 50 \text{ mm}$.

7.4 Kompensierbare Fehler

Neben dem bis hierher erwähnten Fehler treten durch Parameterunsicherheiten des Kameramodells und des Modells des Stereosystems Fehler auf, die aufgrund der weiteren Datenverarbeitung keinen Einfluss auf das Messergebnis haben. Hierzu gehören alle Einflüsse,

die über den gesamten Datensatz einen momentan gleichbleibenden Fehler erzeugen. Diese Fehler werden wegen der implizit differentiellen Auswertung der Messwerte, wie in Kapitel 6.5 beschrieben, kompensiert. Die Punkte, die dem Objekt *Ladegut* zugeordnet werden können, werden, wie in Abschnitt 6.5 beschrieben, in das Koordinatensystem des Laderaumes transformiert. Das Koordinatensystem des Laderaumes wird jedoch aus den rekonstruierten Punkten des *Randes* identifiziert. Somit unterliegen sowohl die Punkte des Ladegutes als auch die Punkte des Randes denselben globalen Fehlern. Ein rotatorischer und translativer Versatz, der für alle rekonstruierten Punkte als hinreichend gleich groß angenommen werden kann, wird somit eliminiert.

7.5 Die Ergebnisse

Für die Bestimmung der Messgenauigkeit des Messsystems *FILLED* wurde ein Referenzkörper mit definierter Geometrie, der einen Ladeguthaufen nachbildet, verwendet. Dieses Referenzmodell wurde aus verschiedenen unbekannten Positionen stereoskopisch aufgezeichnet und ausgewertet. So wurde die Stabilität und Genauigkeit der Messung auch aus verschiedenen Perspektiven überprüft und verifiziert. An dieser Stelle sei angemerkt, dass ausschließlich Messungen verwendet wurden, bei denen der gesamte Referenzkörper sichtbar ist. Darüber hinaus wurden keine Perspektiven gewählt, in denen Teile des Ladeguthaufens durch die Ladebordwände verdeckt sind. Für die Verifikation wird das Epipolarsystem in der in Kapitel 3.4.6 beschriebenen Konfiguration verwendet. Die Abbildung 7.6 zeigt den Versuchsaufbau für die Verifikation.

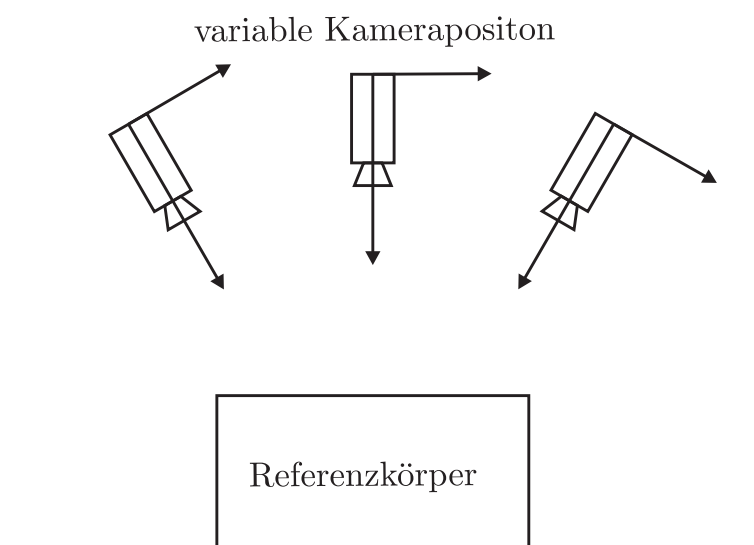


Bild 7.6: Versuchsaufbau für die Verifikation des Messsystems

Die Abbildung 7.7 zeigt die Form des Referenzkörpers.

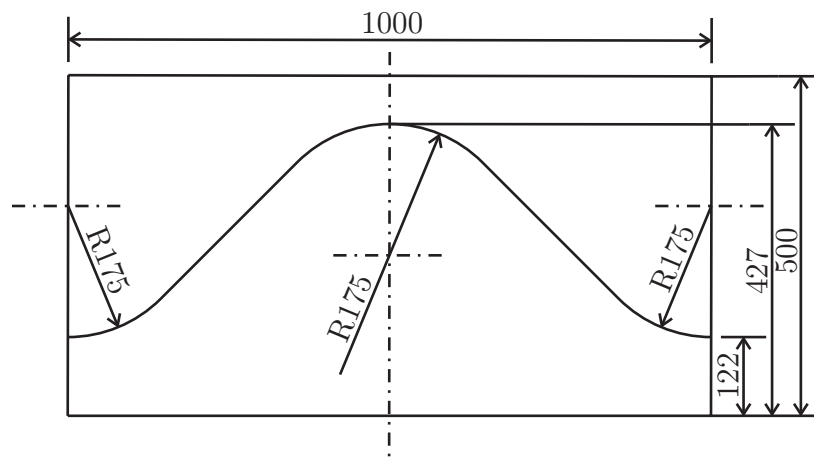


Bild 7.7: Schnitt durch den Referenzkörper

So konnte die lokale Füllhöhe des Referenzkörpers mit einer Standardabweichung von $\sigma_h < 50 \text{ mm}$ bestimmt werden. Die Abbildung 7.8 zeigt die absolute Abweichung der rekonstruierten Oberfläche.

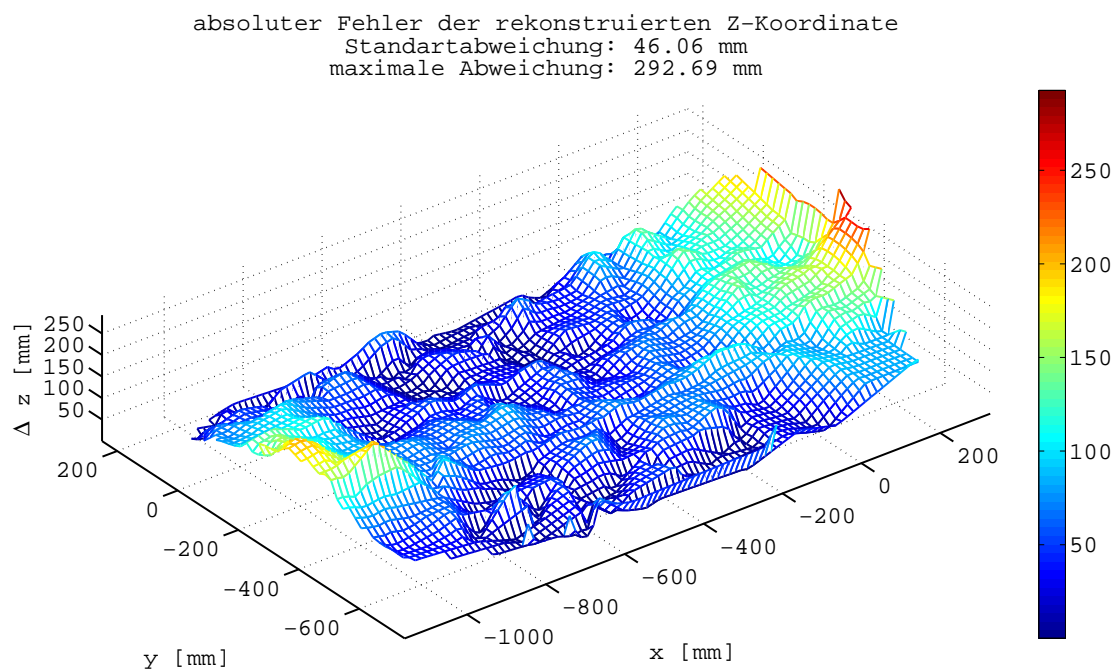


Bild 7.8: Fehler der rekonstruierten Z-Koordinate

Man erkennt deutlich, dass an den Rändern parallel zur Y -Achse die Abweichung der rekonstruierten Oberfläche zum tatsächlichen Objekt größer wird. Dies ist auf den in Abschnitt 7.3 beschriebenen Fehler zurückzuführen, der von ungenauen Verzerrungskoeffizienten herührt. In den Bereichen, um das Zentrum des Datensatzes liegt der maximale Fehler der rekonstruierten Z -Koordinate bei weniger als $\Delta Z = 100 \text{ mm}$. An den Rändern parallel zur Y -Achse wächst die maximale absolute Abweichung der rekonstruierten Z -Koordinate auf bis zu $\Delta Z < 295 \text{ mm}$ an. Je nach geforderter Genauigkeit ist es sinnvoll, den Bereich der weiterzubearbeitenden Daten so einzugrenzen, dass die Anforderungen an die Genauigkeit erfüllt werden. Es wird jedoch deutlich, dass in weiten Bereichen der rekonstruierten Oberfläche der Messfehler geringer ist als $|\Delta Z| = 100 \text{ mm}$ und somit der angestrebten Genauigkeit $\Delta Z_{max} = \pm 100 \text{ mm}$ entspricht.

8 Schlussbemerkungen und Ausblick

In einem Zeitalter, in dem auch die landwirtschaftliche Produktion immer kostengünstiger werden muss, besteht die Forderung nach größeren, schnelleren, leistungsfähigeren und vor allem billigeren Maschinen bzw. Produktionsmethoden. Heutzutage sind für die Verarbeitung von Futtermais bereits selbstfahrende Erntemaschinen mit einer Motorleistung von bis zu 740 kW und einem Gutdurchsatz von bis zu 400 t pro Stunde erhältlich. Eine solche Maschine wiederum füllt einen handelsüblichen Lastkraftwagen in weniger als 5 Minuten. Der Fahrer einer solchen Erntemaschine muss neben der Steuerung der Fahrtrichtung auch die Position des Fahrzeugs zu einem parallel fahrenden Transportfahrzeug, die Ausrichtung des Auswurfkrümmers und die Einstellung der Erntemaschine ausführen, um eine optimale Auslastung der Maschinenleistung sowie der Transportleistung erreichen. Um die Fahrer dieser Erntekombination von einigen Aufgaben während des Erntevorganges zu entlasten, wurde am Institut für Regelungstechnik in Zusammenarbeit mit der DFG und dem Institut für Landmaschinen und Fluidtechnik das Assistenzsystem zur automatisierten Überladung *ASÜL* entwickelt. Die letzte fehlende Komponente für die vollständige Automatisierung des Überladevorganges war ein Messsystem zur Erfassung der Verteilung des Ladegutes im Laderaum der Transporteinheit.

Mit Hilfe eines solchen Messsystems kann nun mehr eine geregelte Befüllung des Laderaumes ermöglicht werden. Neben der automatischen Positionierung des Beladungspunktes im Bereich des Laderaums stellt es eine wesentliche Erleichterung für den Fahrer der Erntemaschine dar, von der Überwachung und Steuerung des Beladungsvorganges entbunden zu werden.

Vor diesem Hintergrund wurde am Institut für Regelungstechnik der Technischen Universität in Braunschweig im Rahmen dieser Arbeit das Messsystem *FILLED* zur Erfassung des Füllstandes einer landwirtschaftlichen Transporteinheit entwickelt.

Das Ziel bestand darin eine Messeinrichtung zu entwickeln, mit deren Hilfe während der Beladung durch einen selbstfahrenden Feldhäcksler der Füllstand des beteiligten Transportmittels erfasst werden kann. Hierbei war insbesondere zu beachten, dass sich die am Überladungsprozess beteiligten Fahrzeuge relativ zu einander in Bewegung befinden. Darüber hinaus war

eine Modifikation des Transportmittels aus ökonomischen Gründen und mit Hinsicht auf die Anwendbarkeit nach Möglichkeit zu vermeiden.

Im Rahmen dieser Arbeit konnte erfolgreich gezeigt werden, dass mit Hilfe eines stereoskopischen Sensorsystems die lokale Füllhöhe des Ladegutes im Rahmen einer Genauigkeit von $\Delta h < \pm 100 \text{ mm}$ gemessen werden kann. Somit ist es gelungen, ein Messsystem zu schaffen, dass es ermöglicht, die Verteilung von gehäckseltem Mais im Laderaum der Transporteinheit mit Bezug auf den Beladungsprozess sowie die Stellgenauigkeit des Beladungspunktes hinreichend genau zu erfassen. Besonderes Merkmal des Systems *FILLED* ist, dass für den momentanen Entwicklungsstand nur die Ränder des Laderaumes farblich markiert werden müssen, somit nur eine minimale Modifizierung des Transportmittels erforderlich ist.

Für eine Erweiterung der Einsetzbarkeit des Messsystems *FILLED* wäre es erstrebenswert, das System dahingehend zu erweitern, dass auf jegliche Modifikation des Transportfahrzeuges verzichtet werden kann. Besonders mit Hinsicht auf eine eventuelle Produktentwicklung gewinnt dieser Schritt an Bedeutung, da dann auf den Einsatz von Farbkameras zugunsten der wesentlich kostengünstigeren Grauwert-Kameras verzichtet werden könnte. Dieses würde jedoch voraussetzen, dass *FILLED* die Ränder des Laderaumes auch ohne farbliche Markierung erkennen kann. Im weiteren Umfeld dieser Arbeit konnte gezeigt werden, dass es grundsätzlich möglich ist, die nicht markierten Kanten des Laderaumes zu erkennen (Graefe/Jansen [4]). Vom Einsatz der hier erprobten Methoden wurde jedoch aufgrund des zu großen Berechnungsaufwandes abgesehen. Darüber hinaus wird zum jetzigen Entwicklungsstand noch von einer rechteckigen Öffnung des Laderaumes und dessen Oberseite ausgegangen. Diese Annahme kann jedoch nicht für die Allgemeinheit der landwirtschaftlichen Transportmittel getroffen werden. Somit wäre es sinnvoll, den Funktionsrahmen von *FILLED* dahingehend zu erweitern, dass auch für Transportanhänger die nicht den oben angeführten Anforderungen genügen, der Füllstand mit den in Kapitel 6.2 beschriebenen Größen gemessen werden kann. Im Weiteren ist die Latenz von ca. 2 Sekunden zwischen Bildeinzug und Messwertausgabe als Nachteil zu nennen. Obwohl bereits großer Wert auf die Minimierung der Zykluszeit des Messsystems gelegt wurde, besteht noch die Möglichkeit durch softwaretechnische Optimierung der Applikation, sowie weitere mathematische Optimierung der verwendeten Algorithmen eine Steigerung der Berechnungsgeschwindigkeit zu erreichen.

Anhand der visuellen Eigenschaften des gehäckselten Mais ist es denkbar, das System *FILLED* auch bei anderen Schüttgütern mit ähnlicher Erscheinung anzuwenden. Es kann angenommen werden, dass durch Anpassung der Applikationsparameter *FILLED* auf weitere Schüttgüter wie zum Beispiel Kies, gehäckseltes Holz, Rüben, o.ä. angewendet werden kann.

Nach erfolgreicher Prüfung der Verwendbarkeit von *FILLED* für andere Güter erscheint die Anbindung an eine Parameterdatenbank sinnvoll, die ein kurzfristiges Anpassen der Applikation auf das momentan vorliegende Material ermöglicht. Ferner wäre die automatische Adaption der Systemparameter ohne den Eingriff des Bedieners erstrebenswert. Dies würde die Entwicklung eines Systems zur Erkennung bzw. Klassifikation der vorliegenden Materialien bedürfen. Vom jetzigen Entwicklungsstand aus gesehen, könnte es hier zielführend sein, die Dynamik der sichtbaren Texturen zu bewerten, um eine Differenzierung der “Feinheit” der Textur zu ermöglichen.

Mit Hinsicht auf den Einsatz auf Landmaschinen oder in ähnlich rauem Umfeld wäre es ratsam, die bis hierher als konstant angenommenen Modellparameter des Stereosystems in regelmäßigen Abständen durch geeignete Verfahren zu prüfen und nötigenfalls das Stereosystem neu zu kalibrieren. Hierfür erscheinen Methoden wie der *Eight-Point Algorithmus*, nach Torr [12] geeignet. Erste Untersuchungen im Rahmen der Entwicklung von *FILLED* konnten die Leistungsfähigkeit des *Eight-Point Algorithmus* nachvollziehen. Eine hinreichende Stabilität und Robustheit für den praxisnahen Einsatz dieses Verfahrens konnte jedoch im Umfeld dieser Arbeit nicht festgestellt werden.

Abschließend bleibt zu erwähnen, dass im Funktionsrahmen von *FILLED* weitere Informationen erzeugt werden, die für die Automatisierung des Überladungsprozesses genutzt werden könnten. Es ist naheliegend, das durch *FILLED* identifizierte Koordinatensystem des Laderaumes für die Bestimmung der Relativposition des Transportmittels zum Häcksler zu verwenden. Nach Graefe/Neermann [5] ist es durch einfache geometrische Analyse der in den Abschnitten 4 und 6 gefundenen Informationen über die Lage der Ränder des Laderaumes möglich, den Beladungspunkt dem Laderaum geregelt nachzuführen. Dies würde den GPS-gerätetechnischen Aufwand, der für das Relativpositionierungssystem *GNATTI* ersetzen, welches bis jetzt im Rahmen von *ASÜL* eingesetzt wird. Somit erscheint es in jedem Fall lohnenswert, das System *FILLED* hinsichtlich weiterer Funktionsintegration auf Basis des zur Verfügung stehenden Daten- und Funktionsrahmen hin zu prüfen.

A Projektive Geometrie

Die projektive Geometrie ist eine Erweiterung der *euklidischen Geometrie* und entsteht aus der Erweiterung des ursprünglich betrachteten Raumes um eine Dimension. Das bedeutet im Prinzip, es werden Punkte als Geraden und Geraden als Ebenen betrachtet. Dies hat den Vorteil, dass Spezialfälle, die in der euklidischen Geometrie gesondert behandelt werden müssen (bspw. Parallelität, Punkte im Unendlichen), in der projektiven Geometrie wie alle anderen Fälle auch behandelt werden können. Da die projektive Geometrie direkt mit der Projektion von 3D-Punkten auf Bildebenen in Zusammenhang steht, bildet sie die Grundlage für das Themengebiet Stereo Vision. Die projektive Ebene \mathcal{P}^2 ist definiert als

$$\mathcal{P}^2 = \mathbb{R}^3 - (0, 0, 0)^T \tag{A.1}$$

A.1 Geraden

Betrachten wir eine Gerade in der Ebene mit der Gleichung $ax + by + c = 0$, die auch alternativ durch den Vektor $(a, b, c)^T$ dargestellt werden kann. Für jeden Faktor $k \neq 0$ gilt, dass die Gleichungen $ax + by + c = 0$ und $(ka)x + (kb)y + (kc) = 0$ gleich sind. Daher repräsentieren die Vektoren $(a, b, c)^T$ und $k(a, b, c)^T$ auch die gleiche Gerade. Vektoren die diese Bedingung erfüllen werden als homogene Vektoren bezeichnet.

A.2 Punkte

Ein Punkt $(x, y)^T$ wird zu einem homogenen Vektor erweitert, indem ihm in der dritten Dimension eine 1 hinzugefügt wird, also $(x, y, 1)^T$. Wie zuvor repräsentieren die Vektoren $(kx, ky, k)^T$ mit verschiedenen Werten für k den Punkt $(x, y)^T$ in \mathbb{R}^2 . Ein beliebiger homogener Vektor mit $x = (x_1, x_2, x_3)^T$ repräsentiert daher den Punkt $(x_1/x_3, x_2/x_3)^T$ in \mathbb{R}^2 .

A.3 Beziehung von Geraden und Punkten

Ein Punkt x liegt genau dann auf einer Geraden l wenn gilt

$$x^T l = 0, \quad (\text{A.2})$$

wobei x, l durch *homogenen Koordinaten* beschrieben sind. Der Schnittpunkt x zweier Geraden l, l' kann durch deren Kreuzprodukt bestimmt werden

$$x = l \times l' \quad (\text{A.3})$$

Analog kann die Gerade l , die durch die beiden Punkte x, x' läuft, durch das Kreuzprodukt der Punkte ermittelt werden.

$$l = x \times x' \quad (\text{A.4})$$

A.4 Dualitätsprinzip

Das Dualitätsprinzip besagt, dass die Rollen von Punkten und Geraden getauscht werden können. D.h. alles, was auf Punkte zutrifft, trifft auch auf Geraden zu, bzw. umgekehrt. Das bedeutet ebenso, dass sobald etwas für Punkte bewiesen wurde dieses ebenso auf Geraden übertragen werden kann, bzw. umgekehrt. Ein Beispiel dieser Dualität ist der Schnittpunkt zweier Geraden, der entsprechend ist zu der Geraden durch zwei Punkte. Bei der Dualität müssen die Begrifflichkeiten ebenfalls übersetzt werden.

A.5 Punkte im Unendlichen

Homogene Vektoren (x_1, x_2, x_3) mit $x_3 \neq 0$ repräsentieren endliche Punkte in \mathbb{R}^2 . Alle Vektoren, die nicht in \mathbb{R}^2 , aber in \mathcal{P}^2 liegen, sind homogene Vektoren mit $x_3 = 0$, sogenannte *uneigentliche* oder *Fernpunkte* (ideal points). Die Menge aller Fernpunkte kann beschrieben werden mit $(x_1, x_2, 0)^T$. Alle diese Punkte liegen auf einer Geraden $l_1 = (0, 0, 1)^T$, die sogenannte *Ferngerade*. Mittels homogener Koordinaten kann der Schnittpunkt x zweier paralleler Geraden $l = (a, b, c)^T$ und $l' = (a, b, c')^T$ bestimmt werden als

$$x = l' \times l = (b, -a, 0)^T. \quad (\text{A.5})$$

Das bedeutet, dass die Geraden sich im Unendlichen treffen. $(b, -a, 0)^T$ gibt die Richtung der beiden parallelverlaufenden Geraden an.

A.6 Modell der projektiven Ebene

Ein Modell der projektiven Ebene \mathcal{P}^2 kann dargestellt werden wie in Abb. A.1 gezeigt. Ein Punkt x auf der Ebene $x_3 = 1$ repräsentiert eine Gerade durch den Ursprung. Eine Gerade l repräsentiert eine Ebene durch den Ursprung, die die Ebene $x_3 = 1$ in l schneidet. Ferngeraden verlaufen parallel zu $x_3 = 1$, sie liegen zusammen mit den Fernpunkten auf der Ebene $x_3 = 0$.

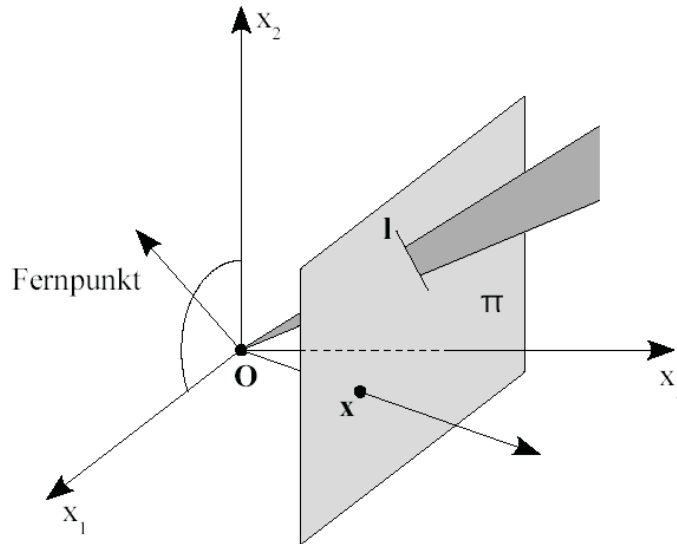


Bild A.1: Ein Modell der projektiven Ebene [24]

A.7 Projektive Transformation

Eine projektive Transformation ist eine nicht-singuläre 3×3 Matrix H mit

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (\text{A.6})$$

Diese Überführung durch die Matrix hat die Eigenschaft, dass Linien erhalten bleiben. Generell entspricht eine projektive Transformation einer Überführung von Punkten einer Ebene im Raum auf eine andere Ebene, die beliebig orientiert sein kann. Abbildung A.2 zeigt solch einen Fall.

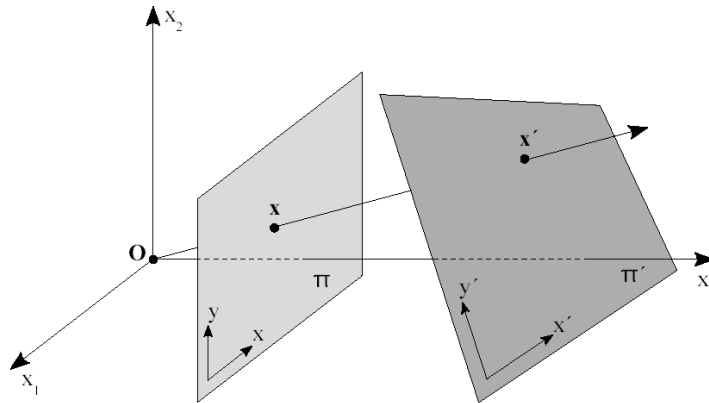


Bild A.2: Die projektive Transformation eines Punktes [24]

B Transformationen im dreidimensionalen Raum

Eine Transformation im dreidimensionalen Raum \mathbb{R}^3 besteht aus einer Rotation (Drehung), einer Translation (Verschiebung) und einer Skalierung.

B.1 Rotation

Dreidimensionale Rotationen können durch eine 3×3 Matrix repräsentiert werden. Mit

$$P_1 = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \cdot P_2 \quad (\text{B.1})$$

wobei $P_1 = (x_1, y_1, z_1)^T$ den rotierten 3D-Punkt und $P_2 = (x_2, y_2, z_2)^T$ den originalen 3D-Punkt darstellt. Im \mathbb{R}^3 ist eine beliebige Rotation geometrisch einfach durch eine Rotation um die drei Koordinatenachsen x, y und z vorzustellen. Im folgenden stellt $R_x(\theta_x)$ eine Rotationsmatrix um die x-Achse dar. Respektive sind $R_y(\theta_y)$ und $R_z(\theta_z)$ zu verstehen.

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 \cos \theta_x & -\sin \theta_x & 0 \\ 0 \sin \theta_x & \cos \theta_x & 0 \end{pmatrix} \quad (\text{B.2})$$

$$R_y = \begin{pmatrix} 1 & 0 & 0 \\ 0 \cos \theta_y & -\sin \theta_y & 0 \\ 0 \sin \theta_y & \cos \theta_y & 1 \end{pmatrix} \quad (\text{B.3})$$

$$R_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 \cos \theta_z & -\sin \theta_z & 0 \\ 0 \sin \theta_z & \cos \theta_z & 1 \end{pmatrix}. \quad (\text{B.4})$$

Wobei θ_x der Winkel der Rotation um die x-Achse, θ_y der Winkel der Rotation um die y-Achse und θ_z der Winkel der Rotation um die z-Achse ist. Bei einer beliebigen Rotation im Raum spielt die Rotationsreihenfolge um die einzelnen Achsen eine wichtige Rolle. Die Rotationsmatrix, die sich aus den elementaren Rotationsmatrizen um die einzelnen Achsen in der Reihenfolge - Rotation um z, Rotation um y und zuletzt eine Rotation um die x-Achse ergibt folgt als Produkt von Matrizen: $R = R_x() \cdot R_y() \cdot R_z()$

$$R = \begin{pmatrix} \cos \theta_y \cos \theta_z & -\cos \theta_y \sin \theta_z & -\sin \theta_y \\ -\sin \theta_x \sin \theta_y \cos \theta_z + \cos \theta_x \sin \theta_z & \sin \theta_x \sin \theta_y \sin \theta_z + \cos \theta_x \cos \theta_z & -\sin \theta_x \cos \theta_y \\ \cos \theta_x \sin \theta_y \cos \theta_z + \sin \theta_x \sin \theta_z & -\cos \theta_x \sin \theta_y \sin \theta_z + \sin \theta_x \cos \theta_z & \cos \theta_x \cos \theta_y \end{pmatrix}. \quad (\text{B.5})$$

Unbedingt zu beachten ist, dass die Rotation nicht kommutativ ist, d.h. $R_x R_y \cdot R_z \neq R_z R_y R_x$. Wichtige Eigenschaften der 3×3 Rotationsmatrix R sind:

- $R^T = R^{-1}$ (die Transponierte der Matrix ist gleich dem Inversen)
- $R \cdot R^T = I$ (wobei I die Einheitsmatrix ist)
- $\text{rank}(R) = 3$
- die Spalten und Zeilen aus R bilden eine orthonormale Basis

B.2 Translation

Eine Translation eines Punktes im \mathbb{R}^3 wird durch einen 1×3 Vektor $\vec{t} = (t_x, t_y, t_z)^T$ durch $P_1 = P_2 + \vec{t}$ berechnet, wobei $P_1 = (x_1, y_1, z_1)^T$ der verschobene und $P_2 = (x_2, y_2, z_2)^T$ der originale Punkt ist.

B.3 Skalierung

Eine Skalierung wird im \mathbb{R}^3 durch eine 3×3 Diagonalmatrix $S = \text{diag}(s_x, s_y, s_z)$ mit $P_1 = S \cdot P_2$ durchgeführt.

B.4 Allgemeine Transformation

Eine beliebige Transformation kann man durch eine Kombination aus Rotation, Translation und Skalierung darstellen:

$$P_1 = S \cdot R \cdot P_2 + t = \begin{pmatrix} s_x r_{11} & r_{12} & r_{13} \\ r_{21} & s_y r_{22} & r_{23} \\ r_{31} & r_{32} & s_z r_{33} \end{pmatrix} \cdot P_2 + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (\text{B.6})$$

C Levenberg-Marquardt-Algorithmus

Der Levenberg-Marquardt-Algorithmus ist ein Verfahren, mit dem nichtlineare Kleinst-Quadrat-Probleme gelöst werden, die vor allem in nichtlinearen Ausgleichsrechnungen auftreten.

Ein Kleinst-Quadrat-Problem ist von der Form

$$\min_x \sum_{j=1}^m r_j^2(x) \quad (\text{C.1})$$

mit $x \in \mathbb{R}^n$. Die Funktion $r_j(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ ist zweimal stetig differenzierbar und werden Residuen genannt. Ein solches Problem muss gelöst werden, wenn beispielsweise ein mathematisches Modell zur durchgeführten Messung gesucht wird. Das Modell ist eine nichtlineare Funktion $f(x; t) : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^l$ mit den unbekannten Parametern x des Modells und den bekannten Parametern $t \in \mathbb{R}^k$ der Messung. Werden in $m > n$ Messungen die Messwerte $y_j \in \mathbb{R}^l$ zu den Parametern t_j für $1 \leq j \leq m$ bestimmt, so wird der Abstand der Messwerte vom Modell $r_j(x) = \|y_j - f(x; t_j)\|^2$. Setzt man diese Residuen in Gleichung C.1 ein, so erhält man direkt das zu lösende Kleinst-Quadrat-Problem

$$\min_x \sum_{j=1}^m \|y_j - f(x; t_j)\|^2. \quad (\text{C.2})$$

Für die Lösung eines Kleinst-Quadrat-Problems mit dem Levenberg-Marquardt-Algorithmus sind die analytischen Residuen und die Jacobimatrix der Residuen erforderlich.

$$J(x) = \begin{pmatrix} \frac{\partial r_1}{\partial x_1} & \dots & \frac{\partial r_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial r_m}{\partial x_1} & \dots & \frac{\partial r_m}{\partial x_n} \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (\text{C.3})$$

Je nach Komplexität der r_j sind die partiellen Ableitungen ebenfalls analytisch gegeben. Schließlich ist der Levenberg-Marquardt-Algorithmus ein iteratives Lösungsverfahren, das einen Startwert x_0 in der Nähe des gesuchten Minimums x benötigt. Die Berechnung einer solchen Näherung ist je nach Problem unterschiedlich. Eine Möglichkeit ist, das Modell $f(x; t)$

linear zu approximieren, so dass es als Matrixprodukt

$$Fx_0 = \begin{pmatrix} \tilde{f}(t_1) \\ \vdots \\ \tilde{f}(t_m) \end{pmatrix} x_0 = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \quad (\text{C.4})$$

gegeben ist. Dabei gilt $\tilde{f}(t_j) \in \mathbb{R}^{l \times n}$ mit $f(x; t_j) \approx \tilde{f}(t_j)x_0$. Das überbestimmte lineare Gleichungssystem kann zum Beispiel mit der Singulärwertzerlegung von F gelöst werden.

D Homogene Matrizen

In der Geometrie erlauben homogene Koordinaten die Darstellung von affinen Transformationen durch eine Multiplikation der Koordinaten mit Matrizen in einfacher Weise.

Rotation, Skalierung, Scherung und Translation eines dreidimensionalen Objektes lassen sich je durch eine 3×3 Matrix beschreiben. Hinderlich ist jedoch, dass die drei erstgenannten Operationen eine Matrizenmultiplikation erfordern, die Translation jedoch eine Addition.

Um eine Translation ebenfalls als Multiplikation berechnen zu können, wird der Raum um eine weitere Dimension erweitert. Eine Translation im dreidimensionalen Raum lässt sich nun durch eine Matrizenmultiplikation mit einer 4×4 Matrix beschreiben.

Die Transformation von kartesischen Koordinaten nach homogenen Koordinaten erfolgt durch:

$$(x, y, z)^T \leftarrow (x, y, z, 1)^T \quad (\text{D.1})$$

Die Abbildung eines Punktes $P_{x,y,z}$ von einem Koordinatensystem in ein anderes geschieht durch die Multiplikation mit der homogenen Matrix M .

$$(x', y', z', w')^T = M_{4 \times 4} \cdot (x, y, z, w)^T \quad (\text{D.2})$$

Aufgrund der Assoziativität von Matrizenmultiplikationen können mehrere aufeinanderfolgende Multiplikationen zu einer einzigen Gesamtmatrix zusammengefasst werden. Die fehlende Kommutivität bedingt, dass die Reihenfolge der Objektmanipulationen von Bedeutung ist. Anschaulich lässt sich dies so interpretieren, dass es einen Unterschied macht, ob in einem lokalen Bezugssystem ein Objekt zuerst entlang einer Achse verschoben und dann im neuen lokalen Bezugssystem rotiert wird, oder ob ein Objekt zuerst rotiert und dann das Objekt in der gleichen Richtung des nun mitrotierten lokalen Bezugssystems in eine neue Richtung in Bezug auf Weltkoordinaten verschoben wird.

Eine Rücktransformation von homogenen Koordinaten in kartesische Koordinaten erlaubt es, auch die perspektivische Abbildung durch eine Matrix zu beschreiben:

$$(x', y', z', w')^T \rightarrow \left(\frac{x'}{w'}, \frac{y'}{w'}, \frac{z'}{w'} \right)^T = (x'', y'', z'')^T \quad (\text{D.3})$$

Die Koordinaten x'' und y'' können bereits als Bildschirmkoordinaten verwendet werden. z'' gibt die Distanz des transformierten Punktes vom virtuellen Bildschirm an und wird im Z-Buffer (Tiefenpuffer) gespeichert. Eine perspektivische Abbildung führt dazu, dass die Gesamtmatrix singulär wird und nicht mehr invertiert werden kann.

E FILLED Systemkonfiguration

Einstellung	Wert	Einheit
CPU	2,4 (32Bit)	GHz
Arbeitsspeicher	1024	MByte
Kameras	$2 \times$ IDS Pacific 9820	k.A.
Framegrabber	$2 \times$ IDS Barracuda CL	k.A.
Bildgröße	640×480	Pixel
Farbtiefe	RGB, 3Byte pro Pixel	Byte
Brennweite f der Objektive	6	mm
Bildqualitäts-Regelerparameter	auf den Blendenwert abgestimmt	k.A.
Bildqualitätssollwert Q_{soll}	206	k.A.
Bildqualitätsparameter a	1	k.A.
Bildqualitätsparameter b	0, 3	k.A.
Bildqualitätsparameter c	15	k.A.
Bildqualitätsparameter $\frac{H_{255}}{H_{255,max}}$	0.05	k.A.
ρ -Auflösung des DeHough-Raumes	255 diskrete Schritte	Pixel
φ -Auflösung des DeHough-Raumes	0.1π	k.A.
DeHough Betriebsmodus	konstante Winkelauflösung	k.A.
Kernelgröße FFT	32×32	Pixel
Schwellwert T_1	$0.75 \cdot f(u, v)_{max}$	k.A.
Anzahl verwendeter Punktepaare	200	Punkte
geforderter Korrelationskoeffizient c_{min}	0.8	k.A.
$XPand$ Radius r	50	mm
Auflösung des Ergebnisdatensatzes	50×50	Punkte

Tabelle E.1: Parameter der erprobten Systemkonfiguration

Literatur

- [1] Fabian Graefe, Raul Queiroz Feitosa “*FILLED Video data based fill level detection of agricultural bulk freight*, ICINCO 2005
- [2] Fabian Graefe “*Assistenzsystem zur Überladung landwirtschaftlicher Güter - Sensorik und Fahrzeugmodelle -*”; DFG Bericht SCHU 977/5
- [3] Cornelia Weltzien “*Assistenzsystem zur Überladung landwirtschaftlicher Güter landwirtschaftlicher Güter - Beladungssteuerung -*”; DFG Bericht HA 1100/17-1
- [4] Fabian Graefe, Ronald Jansen “*Gesteuerte Kantendetektion in Grauwertbildern*”, 2006
- [5] Fabian Graefe, Benjamin Neermann “*Optische Messung des Beladungspunktes eines selbstfahrenden Feldhäckslers*”, 2005
- [6] Danail Stoyanov *A Camera Calibration Toolbox for Matlab*, Royal Society/Wolfson Foundation Medical Image Computing Laboratory, Imperial College London, United Kingdom
- [7] Roger Y. Tsai. *A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses*. IEEE Journal of Robotics and Automation, 1987.
- [8] *Open Source Computer Vision library*,
<http://www.intel.com/research/mrl/research/opencv/index.htm>
- [9] Canny, J. (Nov 1986), “A Computational Approach to Edge Detection”
IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 8, No 6.
- [10] Emanuele Trucco, A. V. (1998)
“Introductory Techniques for 3-D Computer Vision”
Prentice Hall (25. März 1998)
- [11] Jon Orwant, Jarkko Hietaniemi, J. M. D. (April 2000)
“Algorithmen mit Perl”
ISBN 3-89721-141-6, 1. Auflage.
- [12] P.H.S. Torr, A. F. (2004)
“Invariant Fitting of Two View Geometry or In Defiance of the eight point algorithm”
IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(5), pages 648-651
- [13] Will Schroeder, Ken Martin, B. L. (2003)
VTk Toolkit 3rd Edition An Object-Oriented Approach To 3D Graphics,
ISBN 1-930934-12-2, Kitware.

- [14] David A Forsyth, Jean Ponce (2003)
Computer Vision - A Modern Approach,
ISBN 0-13-085198-1, Prentice Hall
- [15] 1 Bruno Caprile and Vincent Torre,
“Using vanishing points for camera calibration”.
The International Journal of Computer Vision,
4(2):127-140, March 1990.
- [16] David A. Forsyth and Jean Ponce,
“Computer Vision: A Modern Approach”,
chapter 1. Prentice Hall, 2002.
- [17] Prof. Dr.-Ing Wolfgang Förstner,
“Photogrammetrie I”,
Universität Bonn, April 2002.
- [18] Richard Hartley,
“Self-calibration from multiple views with a rotating camera”,
Springer-Verlag, 1994.
- [19] Intel Corporation, “Open Source Computer Vision Library”,
Reference Manual, 2001.
- [20] Zhengyou Zhang,
“A Flexible New Technique for Camera Calibration”,
Microsoft Research, December 1998.
- [21] Zhengyou Zhang,
“Flexible Camera Calibration By Viewing a Plane From Unknown Orientations”,
Microsoft Research, 1999.
- [22] Zhengyou Zhang
“Camera Calibration With One-Dimensional Objects”,
Microsoft Research, August 2002.
- [23] Wikipedia
http://de.wikipedia.org/wiki/Projektive_Geometrie, 2005.
- [24] Richard Hartley; Andrew Zisserman,
“Multiple View Geometry”,
Cambridge University Press, 2000.
- [25] Dipl.-Ing. Robert Godding,
“Geometric calibration and orientation of digital imaging systems”,
AICON 3D Systems GmbH.
- [26] Fraunhofer-Institut für Autonome Intelligente Systeme,
„KURT3D“,
<http://ais.gmd.de/ARC/kurt3D/>, April 2003.

- [27] Prof. Dr. Xiaoyi Jiang,
“3D-Model Reconstruction using Vanishing Points”, 2004
- [28] Open Source Computer Vision library,
<http://www.intel.com/research/mrl/research/opencv/index.htm>
- [29] Prof. Dr. X. Jiang,
“Bildanalyse”,
Universität Münster, 2005
- [30] C. Harris; M. Stephens,
“A Combined Corner and Edge Detector”,
4th Alvey Vision Conf. Manchester, S. 189-192, 1988
- [31] H. P. Moravec,
“Towards automatic visual obstacle avoidance”,
Fifth International Joint Conference On Artificial Intelligence, S. 584, 1977
- [32] J. Freixenet and X. Munoz and J. Marti and X. Llado,
“Colour Texture Segmentation by Region-Boundary Cooperation”,
European Conference on Computer Vision, Vol II: 250–261, 2004,
<http://springerlink.metapress.com/openurl.asp?genre=article&issn=0302-9743&volume=3022&spage=250>
- [33] B. S. Manjunath and T. Simchony and R. Chellappa, “Stochastic and Deterministic Networks for Texture Segmentation”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1990, Vol 38-6, p. 1039
- [34] T. Brox and M. Rousson and R. Deriche and J. Weickert,
“Unsupervised Segmentation Incorporating Colour, Texture, and Motion”,
INRIA, 2003, <http://www.inria.fr/rrrt/rr-4760.html>
- [35] W. C. K. Wong and A. C. S. Chung,
“Bayesian Image Segmentation Using Local Iso-Intensity Structural Orientation”,
IEEE Transactions on Image Processing, 2005, Vol. 14-10, 21.12.2005, pp. 1512-1523,
<http://dx.doi.org/10.1109/TIP.2005.852199>
- [36] A. Materka and M. Strzelecki,
Texture Analysis Methods A Review,
Technical University of Lodz, 1998, COST B11 report
- [37] M. Tucur and A. K. Jain,
“Texture analysis”,
Handbook of Pattern Recognition and Computer Vision, 1993, World Scientific, Singapore,
chapter 2, pp. 235-276
editor: C. H. Chen and L. F. Pau and P. S. P. Wang
- [38] R. M. Haralick,
“Statistical and Structural Approaches to Texture”,
Proc. IEEE
1979, may, Vol. 67-5 pp. 786-804

- [39] B. Julesz,
“Experiments in the Visual Perception of texture”, *Scientific American*, Apr.1975, Vol. 232-4, pp. 34-43
<http://iris.usc.edu/Vision-Notes/bibliography/twod315.html#TT17966>
- [40] “Issue on contouring including list of publicly available mapping programs.” *Geobyte* June 1992
- [41] R. Banks
“Contouring Algorithms”, *Geobyte*, Oct. 91
- [42] Smith, W. H. F, and P. Wessel
“Gridding with continuous curvature splines in tension”, *Geophysics*, 1990, Vol 3, 293-305.
- [43] Briggs, I.C.
“Machine Contouring using minimum curvature” *Geophysics*, 1974, Vol 39, 39-48.
- [44] Crain, I.K.
“Computer Interpolation and contouring of 2 dimensional data. A review:”, *Geoexplor* 1970 Vol 8 p 71-86.
- [45] Crain, I.K. and Bhattchauiya, B. K.
“Treatment of non-equispaced, 2 dimensional data with a digital computer”, *Geoexplor* 1967, Vol 5 p 173-194.
- [46] Barnhill, R.E. and Nilesen, G.M.
“Surfaces (the foreward of the special issue) *Rocky Mount J. Math*” 1984, 14, 1-3;
- [47] Inoue, H.
“A least-squares smooth fitting for irregularly spaced data: Finite Element approach using the cubic B spline basis:”, *Geophysics*, 1986, Vol 51, 2051-2066.
- [48] Lancaster, P. QNE wqlkquwkqw, K.
“Curve and surface fitting”, *Academic Press Inc.*, 1986
- [49] Press, W. H., Flannery, B.P., Teukolsky, S. A., and Vetterling, W.T.
“Numerical Recipies”, *Cambridge Univ. Press.*, 1986
- [50] Olea, R.A.
“ Optimal contour mapping using universal kriging”, *J. Geophys. Res.*, 1974, Vol 79, 696-702.
- [51] Belongie, Serge
“Rodrigues’ Rotation Formula.” *From MathWorld–A Wolfram Web Resource, created by Eric W. Weisstein.* <http://mathworld.wolfram.com/RodriguesRotationFormula.html>
- [52] Couthard, John
“QUIKGRID, 3-D rendering of a surface represented by scattered data points”
<http://www.perspectiveedge.com/references.html>
- [53] <http://www.fe-lexikon.info/lexikon-g.htm>

[54] *<http://fourier.eng.hmc.edu/e161/lectures/gradient/node2.html>*